

Fri May 24 11:27:35 2002

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model
Run on: May 24, 2002, 10:12:57 : Search time 101.42 Seconds
(without alignments)
516.928 Million cell updates/sec

Title: WALICK-934-95.PEP
Sequence: 1 MKPTSLWMSAGALALALAP.....GMOYVALAGGVSYSGINS 472

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5
Searched: 747574 seqs, 111073796 residues
Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq-032802:*

- 1: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1980.DAT:*
- 2: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1981.DAT:*
- 3: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1982.DAT:*
- 4: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1983.DAT:*
- 5: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1984.DAT:*
- 6: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1985.DAT:*
- 7: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1986.DAT:*
- 8: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1987.DAT:*
- 9: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1988.DAT:*
- 10: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1989.DAT:*
- 11: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1990.DAT:*
- 12: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1991.DAT:*
- 13: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1992.DAT:*
- 14: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1993.DAT:*
- 15: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1994.DAT:*
- 16: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1995.DAT:*
- 17: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1996.DAT:*
- 18: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1997.DAT:*
- 19: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1998.DAT:*
- 20: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1999.DAT:*
- 21: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA2000.DAT:*
- 22: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2478	97.9	579	19 AAW37873	Alcohol and/or ald
2	2153	85.1	579	19 AAW37874	Alcohol and/or ald
3	2145.5	84.8	579	19 AAW37875	Alcohol and/or ald
4	2065.5	81.6	578	19 AAW37875	Sorbitol dehydroge
5	413	16.3	754	21 AAB35987	Amino acid sequenc
6	408	16.1	732	11 AAR05235	ADH complex protei
7	407.5	16.1	738	13 AAR05192	A. alioceles genes me
8	406.5	16.1	738	12 AAR13993	Sorbitol dehydroge
9	274.5	10.8	740	20 AAW95019	Novel human diagno
10	137	5.4	443	22 AAB24430	S. epidermidis ope
11	122.5	4.8	715	22 AAG81738	

12	122.5	4.8	948	22 AAG25909	Novel human diagno
13	122.5	4.8	1510	22 AAG21573	Novel human diagno
14	122.5	4.8	910	22 AAG83007	S. epidermidis ope
15	121.5	4.7	886	22 AAG82914	S. epidermidis sequ
16	119	4.7	886	22 AAG83996	Human protein sequ
17	118.5	4.7	770	22 AAG82701	S. epidermidis ope
18	118	4.7	770	22 AAG80824	Amino acid sequenc
19	115.5	4.5	501	18 AAW10645	HPV6 mutant L1 pro
20	114.5	4.5	501	18 AAW10644	HPV6 mutant L1 pro
21	113.5	4.5	1016	18 AAW33449	Staphylococcus aur
22	113	4.5	824	18 AAW09614	Pathogenic Staphyl
23	111.5	4.4	501	18 AAW26108	HPV11 mutant L1 pr
24	111.5	4.4	501	18 AAW26109	HPV11 mutant L1 pr
25	110.5	4.4	501	18 AAW26107	HPV 6 L1 protein m
26	110.5	4.4	501	21 AAW99981	Streptococcus GX78
27	110.5	4.4	594	12 AAR10005	Human HPV6 L1 prot
28	108.5	4.3	530	22 AAW32080	Human PRO3440 poly
29	108.5	4.3	330	22 AAW12236	HPV6 L1 protein.
30	108	4.3	500	18 AAW24582	Human HPV6 L1 prot
31	107.5	4.2	501	18 AAW32082	Human HPV6 L1 prot
32	107.5	4.2	501	18 AAW15110	Human HPV6 L1 prot
33	107.5	4.2	501	18 AAW26113	HPV11 mutant L1 pr
34	107.5	4.2	501	18 AAW26115	HPV11 mutant L1 pr
35	107.5	4.2	501	18 AAW26116	HPV11 mutant L1 pr
36	107.5	4.2	501	18 AAW26106	Protein G variant.
37	107	4.2	269	11 AAR07006	IgG-binding Strept
38	107	4.2	269	15 AAR53295	S. epidermidis ope
39	106.5	4.2	404	22 AAG82772	Human HPV6 L1 prot
40	106.5	4.2	501	18 AAW15109	Human HPV6 L1 prot
41	106.5	4.2	501	18 AAW15111	Human HPV6 L1 prot
42	106.5	4.2	501	18 AAW10646	HPV11 mutant L1 pr
43	106.5	4.2	593	11 AAR07014	Protein G variant
44	106.5	4.2	2599	21 AAW75098	Neisseria meningit
45	106	4.2	269	10 AAW94788	Protein G variant.

ALIGNMENTS

RESULT 1	AAW37873	standard: Protein; 579 AA.
ID	AAW37873	
AC	AAW37873	
XX	10-AUG-1998	(first entry)
DE	Alcohol and/or aldehyde dehydrogenase A amino acid sequence.	
DT	Alcohol/aldehyde dehydrogenase A enzyme; recombinant organism;	
XX	aldehyde; ketone; carboxylic acid; L-sorbose; D-sorbitol;	
XX	KW 2-keto-L-gulononic acid; L-ascorbic; inhibition.	
KW	Glucanobacter oxidans.	
XX	OS	
XX	Key	Location/Qualifiers
FX	Peptide	1..23
FT	Protein	/note= "signal peptide"
FT		24..579
FT		/note= "mature protein"
FT		
XX	EP832974-A2.	
PN	01-APR-1998.	
PD	11-SEP-1997;	97EP-0115801.
XX	19-SEP-1996;	96EP-0115001.
XX	(HOFF) HOFFMANN LA ROCHE & CO AG F.	
PA	Asakura A, Hoshino T, Ojima S, Shiojoh M, Tomiyama N;	
XX		
PI		
XX		

DR WPI: 1998-195228/18.
N-PSDB: AAV29051.

Recombinant Gluconobacter oxydans alcohol and/or aldehyde dehydrogenase enzyme(s) - useful for converting substrate(s), e.g. L-sorbose or D-sorbitol to 2-keto-L-gulononic acid

Claim 1; Pages 35-37; 59pp; English.

This is the amino acid sequence for the Gluconobacter oxydans alcohol and/or aldehyde dehydrogenase A' enzyme. The enzymes or recombinant organisms can be used to convert suitable substrates to aldehydes, ketones or carboxylic acids, especially L-sorbose or D-sorbitol to 2-keto-L-gulononic acid, which can be converted to L-ascorbic acid by standard procedures. The derivatives of ADH enzymes have desired substrate specificity, higher affinity to a substrate, lower affinity to an inhibitory compound, higher stability against temperature and/or pH and higher catalytic speed.

Sequence 579 AA;

Query Match 97.9%; Score 2478; DB 19; Length 579;
Best Local Similarity 84.9%; Pred. No. 2.2e-210;
Matches 472; Conservative 0; Mismatches 0; Indels 84; Gaps 1;

```

OY 1 MKPTSLMASAGALALAAPAFQVTPYDELAPPAEMISYGOENRHRSPQTQT 60
DB 1 mkptslmasaagalalaaafataytpvtdeallampagewisygngenyhspilqit 60
OY 61 TENVGOLQLVWARGMOPCKQVTPPLIHGVMYLAN----- 95
DB 61 tenvgqlqlvwargmopckqvtpplihgvmylanpgdvlgaidaklgdlwehrrqlpn 120
OY 96 ----- 96
DB 121 iatlnsfgeptirgmalylgtlnvfyvswdnhlvaldtagvtfvdvtrggedmvnsnsgpl 180
OY 97 VANGYIVAGSTCOYSPFCFVSGHDSATGELMRNYFIRPACBEGDETWGNDYEARMGTG 156
DB 181 vangyivagstcocyfscfvgshdsatgelmryfipragsegdetwngndyearmgtg 240
OY 157 AMGOITVDPVTNLVHYGSTAVGPASSTORGPSTGLTGNTNRFAVRPDDGEIYWRHQTLP 216
DB 241 awgiltvdpvtnlvhygstavgpasstortgpgstglgtgntnrfaavrpdgeiylwrhqtlp 300
OY 217 RDNNDECTTEEMAVTNVDQVSTEMEGLOSINPMANAGERVLTGVPCKTGTMOFDAET 276
DB 301 rdnndecttemavtnvdqvstemeqlosinpmanagervltgvpcktgtmofdaet 360
OY 277 GEFLMARDNTYONMIESIDENGITVNEDEALIKELDEYDVCPFTLGGDMPSSALMPDS 336
DB 361 geflmardntyonmiesidengitvnedalikeldevdvcptflggdmpssalmpds 420
OY 337 GIYFPLNNVGYDMMAVDQETSMADYNNYSNVTKLPPGKDMIGRDAADISTGRTLMASVE 396
DB 421 giyfplnnvgydmavdqetfsmadynnysnvtklppgkdmigrdaadistgtrtlmasve 480
OY 397 RAAANYSVIVSTGGVLFNGCTDRYFRALSOETGTTLMQTRLATVASGOAISTEYVGIMOY 456
DB 481 raaanysvivistggvlfngctdryfralsetgttltmqtrlatvasgoaisteeyvgimoy 540
OY 457 VALAGGVSTGSGLNS 472
DB 541 valaggvstsgslns 556

```

RESULT 2
AAW37874
ID AAW37874 standard; Protein; 579 AA.
XX
AC AAW37874;
XX

DT 10-AUG-1998 (first entry)

XX Alcohol and/or aldehyde dehydrogenase A' amino acid sequence.

DE Alcohol/aldehyde dehydrogenase A' enzyme; recombinant organism;
KW 2-keto-L-gulononic acid; L-ascorbic; inhibition.

OS Gluconobacter oxydans.

FT Key Location/Qualifiers

FT Peptide 1..23

FT Protein /note= "signal peptide"

FT /note= "mature protein"

PN EP832974-A2.

PD 01-APR-1998.

PF 11-SEP-1997; 97EP-0115801.

PR 19-SEP-1996; 96EP-0115001.

PA (HOFF) HOFFMANN LA ROCHE & CO AG F.

PI Asakura A, Hoshino T, Ojima S, Shinjoh M, Tomiyama N;

DR WPI: 1998-195228/18.

N-PSDB: AAV29052.

Recombinant Gluconobacter oxydans alcohol and/or aldehyde dehydrogenase enzyme(s) - useful for converting substrate(s), e.g. L-sorbose or D-sorbitol to 2-keto-L-gulononic acid

Claim 1; Pages 38-40; 59pp; English.

This is the amino acid sequence for the Gluconobacter oxydans alcohol and/or aldehyde dehydrogenase A' enzyme. The enzymes or recombinant organisms can be used to convert suitable substrates to aldehydes, ketones or carboxylic acids, especially L-sorbose or D-sorbitol to 2-keto-L-gulononic acid, which can be converted to L-ascorbic acid by standard procedures. The derivatives of ADH enzymes have desired substrate specificity, higher affinity to a substrate, lower affinity to an inhibitory compound, higher stability against temperature and/or pH and higher catalytic speed.

Sequence 579 AA;

```

OY 1 MKPTSLMASAGALALAAPAFQVTPYDELAPPAEMISYGOENRHRSPQTQT 60
DB 1 mkptslmasaagalalaaafataytpvtdeallampagewisygngenyhspilqit 60
OY 61 TENVGOLQLVWARGMOPCKQVTPPLIHGVMYLAN----- 95
DB 61 tenvgqlqlvwargmopckqvtpplihgvmylanpgdvlgaidaklgdlwehrrqlpn 120
OY 96 ----- 96
DB 121 iatlnsfgeptirgmalylgtlnvfyvswdnhlvaldtagvtfvdvtrggedmvnsnsgpl 180
OY 97 VANGYIVAGSTCOYSPFCFVSGHDSATGELMRNYFIRPACBEGDETWGNDYEARMGTG 156
DB 181 vangyivagstcocyfscfvgshdsatgelmryfipragsegdetwngndyearmgtg 240
OY 157 AMGOITVDPVTNLVHYGSTAVGPASSTORGPSTGLTGNTNRFAVRPDDGEIYWRHQTLP 216
DB 241 awgiltvdpvtnlvhygstavgpasstortgpgstglgtgntnrfaavrpdgeiylwrhqtlp 300

```

Query Match 85.1%; Score 2153; DB 19; Length 579;
Best Local Similarity 72.8%; Pred. No. 1.3e-181;
Matches 402; Conservative 32; Mismatches 34; Indels 84; Gaps 1;

```

OY 1 MKPTSLMASAGALALAAPAFQVTPYDELAPPAEMISYGOENRHRSPQTQT 60
DB 1 mkptslmasaagalalaaafataytpvtdeallampagewisygngenyhspilqit 60
OY 61 TENVGOLQLVWARGMOPCKQVTPPLIHGVMYLAN----- 95
DB 61 tenvgqlqlvwargmopckqvtpplihgvmylanpgdvlgaidaklgdlwehrrqlpn 120
OY 96 ----- 96
DB 121 iatlnsfgeptirgmalylgtlnvfyvswdnhlvaldtagvtfvdvtrggedmvnsnsgpl 180
OY 97 VANGYIVAGSTCOYSPFCFVSGHDSATGELMRNYFIRPACBEGDETWGNDYEARMGTG 156
DB 181 vangyivagstcocyfscfvgshdsatgelmryfipragsegdetwngndyearmgtg 240
OY 157 AMGOITVDPVTNLVHYGSTAVGPASSTORGPSTGLTGNTNRFAVRPDDGEIYWRHQTLP 216
DB 241 awgiltvdpvtnlvhygstavgpasstortgpgstglgtgntnrfaavrpdgeiylwrhqtlp 300

```



```

OY 217 RDNDDECTFEKMAVTVNDVQSTENEGLOSINPNAATGERRVLGVPCKTGTMOFPAET 276
DB 301 rdnddectfeemvanvdyqpaadmdyrsinpaatgerrvltgpccktgtmwfdaet 360
OY 277 GEELIARDNTYQNMIESIDENGIVTNEEDALIKELDEYDVCPTLGGDRMPAALNPS 336
DB 361 geflwardtvenllesidengivtvedeskvlteldpydvcplllggridwpsaalnps 420
OY 337 GIVFIPLNVCYDMAVDOETSMOVTNTSVTKRPECKDMIGRIDADISTGRTLSVE 396
DB 421 givfiplnvcmdleavdeffsaldyngslakmapkclvgrlidaistgtrltwae 480
OY 397 RAANYSFVLSSTGGVYFNGCDRFRALSOETGELTMOTRIATVAGQALSYVDNMOY 456
DB 481 reasnypavlstagvflngtdrfralsgetgelwqtrlatvasqavseidvgy 540
OY 457 VAIAGGVSYGS 468
DB 541 vaiaggvltys 552

RESULT 3
AAM37876 standard; Protein: 579 AA.
AAM37876;
AAM37876;
10-ANG-1998 (first entry)
Alcohol and/or aldehyde dehydrogenase B amino acid sequence.
Alcohol/aldehyde dehydrogenase B enzyme; recombinant organism;
aldehyde; ketone; carboxylic acid; L-sorbose; D-sorbitol;
2-keto-L-gulonic acid; L-ascorbic; inhibition.
Gluconobacter oxydans.
OS
XX
XX Key Location/Qualifiers
FH Peptide 1..23
FT /note="signal peptide"
FT 24..579
FT Protein /note="mature protein"

EP832974-A2.
PN
XX 01-APR-1998.
XX 11-SEP-1997; 97EP-0115801.
XX 19-SEP-1996; 96EP-0115001.
XX
XX (HOFF) HOFFMANN IA ROCHE & CO AG F.
XX
XX Asakura A, Hoshino T, Ojima S, Shinjoh M, Tomiyama N;
XX WPI; 1998-195228/18.
XX N-PSDB; AAV29054.
XX
XX Recombinant Gluconobacter oxydans alcohol and/or aldehyde
XX dehydrogenase enzyme(s) - useful for converting substrate(s), e.g.
XX L-sorbose or D-sorbitol to 2-keto-L-gulonic acid
XX
XX Claim 1; Pages 44-46; 59pp; English.
XX
XX This is the amino acid sequence for the Gluconobacter oxydans
XX alcohol and/or aldehyde dehydrogenase B enzyme. The enzymes
XX or recombinant organisms can be used to convert suitable substrates
XX to aldehydes, ketones or carboxylic acids, especially to convert
XX L-sorbose or D-sorbitol to 2-keto-L-gulonic acid, which can be
XX converted to L-ascorbic acid by standard procedures. The derivatives
XX of ADH enzymes have desired substrate specificity, higher affinity
XX to a substrate, lower affinity to an inhibitory compound, higher

```

```

CC stability against temperature and/or pH and higher catalytic speed.
XX
XX Sequence 579 AA;
XX
Query Match 84.8%; Score 2145.5; DB 19; Length 579;
Best Local Similarity 71.6%; Pred. No. 5,8e-181; Indels 85; Gaps 1;
Matches 399; Conservative 34; Mismatches 39;
OY 1 MKPTSLMWSACALALLAARPAFOVPTPTDELANPPAGEMISYGOENRHSPLTQIT 60
DB 1 mpptllrtaavvllltpaafaqvptltdellanppagewlnygrngenytrhspltl 60
OY 61 TENVGOLQIYVWAKMOPKQVYTPLIHDGVMYLAN----- 95
DB 61 adnvgqltqvwakmgaagavtptmldhgvmylanppgdvlgalaqtdliwchrrqlpa 120
OY 96 ----- 95
DB 121 valnaqgdtkrgvalytslyfsswdnhllaldmetqgvfdvergsgedjlsnltgp 180
OY 96 IVANGVIVAGSTCOYSPFGCFVSGHDSATGEBLMNRYTPRAGEGGETMGNDEYEAHMT 155
DB 181 ivangvivaagstcogspygcfisghdsatgeelwnrhifpqgeegdetwgndecarmt 240
OY 156 GAMQITVDVPTNLYHGSTAVGPASSETGRTGCTLXGNTFRVAPDGTGEIYWRQTL 215
DB 241 gvwgqitvdvptnlvhygstavgpasetgrtgpctlxgntfrvavpdgtgeiwrqtl 300
OY 216 PRDNDDECTFEKMAVTVNDVQSTENEGLOSINPNAATGERRVLGVPCKTGTMOFPAE 275
DB 301 prdnddectfeemvanvdyqpaadmdyrsinpaatgerrvltgpccktgtmwfdaet 360
OY 276 TGEELIARDNTYQNMIESIDENGIVTNEEDALIKELDEYDVCPTLGGDRMPAALND 335
DB 361 tgeelwardtvenllesidengivtvedeskvlteldpydvcplllggridwpsaalnps 420
OY 336 SGIVFIPLNVCYDMAVDOETSMOVTNTSVTKRPECKDMIGRIDADISTGRTLSV 395
DB 421 sgivfiplnvcmdleavdeffsaldyngslakmapkclvgrlidaistgtrltwae 480
OY 396 ERAANYSFVLSSTGGVYFNGCDRFRALSOETGELTMOTRIATVAGQALSYVDNMOY 456
DB 481 eraanysfvlstaggvflngtdrfralsgetgelwqtrlatvasqavseidvgy 540
OY 456 YVAIAGGVSYGSGLNS 472
DB 541 yvaiaggvltysgtqlna 552

RESULT 4
AAM37875 standard; Protein: 578 AA.
AAM37875;
AAM37875;
10-ANG-1998 (first entry)
Alcohol and/or aldehyde dehydrogenase A'' amino acid sequence.
Alcohol/aldehyde dehydrogenase A'' enzyme; recombinant organism;
aldehyde; ketone; carboxylic acid; L-sorbose; D-sorbitol;
2-keto-L-gulonic acid; L-ascorbic; inhibition.
Gluconobacter oxydans.
OS
XX
XX Key Location/Qualifiers
FH Peptide 1..23
FT /note="signal peptide"
FT 24..578
FT Protein /note="mature protein"

EP832974-A2.
PN

```

XX 01-APR-1998.
 PD 11-SEP-1997; 97EP-0115801.
 XX 19-SEP-1996; 96EP-0115001.
 XX (HOFF) HOFFMANN LA ROCHE & CO AG F.
 PA Asakura A, Hoshino T, Ojima S, Shinjoh M, Tomiyama N;
 PI WPI: 1998-195228/18.
 DR N-PSDB: AAV29053.
 DR
 XX
 PT Recombinant Gluconobacter oxydans alcohol and/or aldehyde
 PT dehydrogenase enzyme(s) - useful for converting substrate(s), e.g.
 XX L-sorbitol or D-sorbitol to 2-keto-L-gulononic acid
 PS Claim 1; Pages 41-43; 59pp; English.
 CC This is the amino acid sequence for the Gluconobacter oxydans
 CC alcohol and/or aldehyde dehydrogenase A' enzyme. The enzymes
 CC or recombinant organisms can be used to convert suitable substrates
 CC to aldehydes, ketones or carboxylic acids, especially to convert
 CC L-sorbitol or D-sorbitol to 2-keto-L-gulononic acid, which can be
 CC converted to L-ascorbic acid by standard procedures. The derivatives
 CC of AADH enzymes have desired substrate specificity, higher affinity
 CC to a substrate, lower affinity to an inhibitory compound, higher
 CC stability against temperature and/or pH and higher catalytic speed.
 XX
 XX Sequence 578 AA;

Query Match
 Best Local Similarity 81.6%; Score 2065.5; DB 19; Length 578;
 Matches 388; Conservative 37; Mismatches 45; Indels 85; Gaps 2;

QY 1 MKPTSLMASGALALALAAPAFQVYPTDELNPAGEMISYGOQENYRHSPLTOIT 60
 DB 1 mktlllgssaalivgtipalag-taltdemlanppagewinygqenyrhsplttgt 59
 QY 61 TENVGLOLVWARGKOPGVQVTPPLIHGVNVLN----- 95
 DB 60 adnvgqdlvwarigmeagkigtvrlphdgvmylanpddvgaalaaatgdliwetrqjpn 119
 QY 96 ----- 1
 DB 120 iatlnsfgeptrgmaljytnyfvswdnhvaldstgqvfvldvdrqgtdmvsnsqpi 179
 QY 97 VANGYIVAGSTCOYSPFCFVSGHDSATGELMRNFIRACBEDDEMNGNDYKRWMTG 156
 DB 180 vanygivaagstcqvspfcfsvghdsatgaelwrnflpraggegcwvndyaaarwmtg 239
 QY 157 AMGOITDYPTNLVHGSTVGPASFTOGCTEGGLYGTNTRFAVRPDTGIVWRHTLP 216
 DB 240 vwgqitdyprvgvlvhygtsavpaeeelqrgtlvgysmytlntrfavrpegtglvwrhtlp 299
 QY 217 RDNDDOCTETEMWNTNVDVOSTEMGLQINPNNAATGERVNLGVCKCTGMQEDAE 276
 DB 300 rdnddgectetemmwnvndvqpsaameglhainpdaatgerivvvgvpcnknglmqidaet 359
 QY 277 GEFLIARDNTYQNMIESIDENGIYTVNEDAIKELDVEYDVCPTEFLGGRDWPSSAALNPDS 336
 DB 360 gefliardntsyqnllsvdgdvlvhnmedlvtelevayelcplflggridwpssaalnpt 419
 QY 337 GIEFILNVVCYDMAVDEFTSMQVNTSVNLTLPKGMIGRIDAIDISGRITLMSVE 396
 DB 420 gylfilnvvcydmavdeftsmqvntsvnltlpkgmigridaidistgritlwaae 479
 QY 397 RAAATSPVLSTGGVLNFGCTDRFRALSOERGETIMQRLATVASSGQAISEYEDGQY 456
 DB 480 ryasnypavlstggvlnfgctdrfralsergetlwtqrlatlvassgqaiselyedgqy 539

QY 457 VALAGGVSYSGSLN 471
 DB 540 valgrggtysynshn 554

RESULT 5
 ID AAB35987
 XX AAB35987 standard; Protein; 754 AA.
 AC AAB35987;
 XX
 DT 01-MAR-2001 (first entry)
 DE Sorbitol dehydrogenase subunit 1 amino acid sequence.
 XX
 KW Sorbitol dehydrogenase; SDH; Gluconobacter oxydans; D-sorbitol;
 KW L-sorbitose production; 2-keto-L-gulononic acid.
 OS Gluconobacter oxydans.
 PN WO200065066-A1.
 PN
 PD 02-NOV-2000.
 PD
 PF 23-APR-1999; 99WO-IB00736.
 PF
 PR 23-APR-1999; 99WO-IB00736.
 XX
 PA (CHOI/) CHOI E.
 PA (RHEE/) RHEE S.
 PA (LEE/) LEE E.
 XX
 PI Choi E, Rhee S, Lee E;
 DR WPI: 2000-687351/67.
 DR N-PSDB: AAC83153.

Novel membrane-bound sorbitol dehydrogenase nucleic acid molecule
 PT isolated from Gluconobacter suboxydans useful for the fermentative
 PT production of 2-keto-L-gulononic acid and L-sorbitol from D-sorbitol
 PS
 PS Claim 1; Fig 8; 96pp; English.

This invention relates to an isolated membrane-bound sorbitol
 CC dehydrogenase (SDH) from Gluconobacter oxydans. The invention includes
 CC polynucleotides (AAC83153 - AAC83154) encoding the 3 subunits of SDH
 CC (AAB35987 - AAB35989). Also included in the invention are two
 CC polynucleotide sequences AAC83156 and AAC83157 which contain the subunit
 CC useful for producing L-sorbitol from D-sorbitol and for increasing the
 CC production of 2-keto-L-gulononic acid by transforming a host cell,
 CC especially Gluconobacter with the DNA and selecting the transformed host
 CC genes) encoding fragments of SDH sequences (fragments of the SDH subunit
 CC sequences are not given separately in the specification but are included
 CC in sequences AAC83156 and AAC83157.

Sequence 754 AA;

Query Match
 Best Local Similarity 16.3%; Score 413; DB 21; Length 754;
 Matches 143; Conservative 72; Mismatches 222; Indels 164; Gaps 15;

QY 14 LALLAAPFAOVTPV-----TDELNPAGEMISYGOQENYRHSPLTOITENYVQ 66
 DB 17 lgcnaalafcatlspvalaetfainadqhgdmmsygrtysqrgysplqdkdhaan 76
 QY 67 LQLVWARGMQPGKVOY-TPLIHGVNVLN-----ANIVAN 99
 DB 77 lklawhyldltnngvgqetplivdgvmyaltwnskmkaalaaatgklwysdpkvpynladr 136
 QY 100 G----- 100

```

Db 137 gccdtvnrqaaayngkvfygftdgrllalaadkqglwswylyvpkaqjghnqrsylvdga 196
OY 101 -----VIVAGSTQCYSPFGCFVSGHDSATGELMNRNFIIPRAGECD----- 142
OY 197 prlqkqgvlignggaefgarg-fvlaydaetgkmwrficvprpdkkpdgaadavlmek 255
OY 143 --ETWGDYDARMMTG--AMGQITVDPTNLVHGSTAVGPASEHQRPGTILGTNT 197
OY 256 aypfwgqgawkgqg99gglwdslllydprdllylgvngsgspwkytriseqgnllfgs 315
OY 198 RFVAVRPTGELIWMHQTLPRDNDQECTFEEMAVTVNDVOPSTEMEGLOSTINPNATGE-R 256
OY 316 lvalpdtgkyvwhfgeqpmqdwqysvqimalampv-----ngemr 358
OY 257 RVLTVGPCKTGMNQPAETGEELMARDINYONMIESIDE-NGIYVNEDEALILKELVEY 315
OY 359 hvlvnap-kngfflyldakqglfsgkpylyenwangldpvtgtrpnydpdalwlnqkpw 417
OY 316 DWCPTEFGSDMPASALNPDGSIYFIPLNVCY-----DMNAVDEETSMGV 362
OY 418 yslpgqlgghmfaamayspckllylpaqvyrfvydpqkgfkahdswnlglmklgl 477
OY 363 YMTSN-----VTKLPQKDMIGRIDALIDISTGRTLSVERAANYSPLSTGGVLENGCT 418
OY 478 lddndpqhkdqkqflkalkgwlvawdpqkaafvdkhgwngjllataqglfqla 537
OY 419 DRYFRALSQETGELMOTRLATVASGOALSYEVDMQYVAI---AG-----GGVSYGS 468
OY 538 ngethaydaltgkdlitfpaqsaliprvtylangkyvavewg9lylplfllgyvarcs 597
OY 469 G 469
OY 598 g 598

```

RESULT 6
AAR05235 standard; protein: 742 AA.

AA05235;
04-AUG-1990 (first entry)

XX DT Amino acid sequence of alcohol dehydrogenase (ADH) encoded by bases
XX DE 1-229 of alcohol dehydrogenase (ADH) gene.
XX KM Alcohol dehydrogenase (ADH) gene; Acetobacter pasteurianus IFO 3191;
XX KW Acetobacter aceti K1006 (FERM-7528); plasmid pAA721.
XX OS Acetobacter aceti strain K1006 (FERM-7528).
XX PN JF02000452-A.
XX PD 05-JAN-1990.
XX PF 30-OCT-1987; 87JP-0273190.
XX PR 30-OCT-1987; 87JP-0273190, JP-075069.
XX PA (QPP) OP CORP (QPPJ).
XX DR WPI: 1990-047990/07.
XX DR N-PSDB; AAG91811.
XX PT Cloning DNA, plasmid and microbe contg. It
XX PT contains alcohol dehydrogenase gene derived from Acetobacter
XX PT aceti K10006 strain
XX PS Disclosure; fig 4; Bpp; Japanese.
XX CC Also new are a recombinant plasmid contg. its encoding DNA, and a
XX CC microorganism transformed with the plasmid.

A DNA fragment was obtd. from A. aceti strain K1006 and was transferred to an ADH-defective strain. The resulting plasmid, pAA721, contg. its gene was inserted into A. pasteurianus strain IFO 3191 by the triparent method. Its gene is useful for improving Acetobacter culturing.

Sequence 742 AA:

Query Match 16.1%; Score 408; DB 11; Length 742;
Best Local Similarity 24.2%; Pred. No. 4,3e-27;
Matches 149; Conservative 78; Mismatches 192; Indels 196; Gaps 22;

```

OY 11 AGALALLAFAFQVRYPTDEL--ANPPAGWISYQONQENRHSPLQITTEVNGQL 68
OY 23 aaalpyaavpraagngntgealhhadhpewlsyrglyseqytspldqinsvngdlk 82
OY 69 LVMARQOPKQV-VTPLIHDGMYL-----ANI----- 96
OY 83 llyytlidnrgaetrlvvdgimyatnmskmealdaatyklwqyqpkvgnladdgc 142
OY 97 ----- 96
OY 143 cdtvnrqaaayngkvfygftdgrllvaadaktgkvwavnltpadaslgkqrsylvdgvtr 202
OY 97 VANGVIVAGSTQCYSPFGC--FVSGHDSATGELMNRNFIIPRAGECD----- 142
OY 203 vakglvllgn--ggaelfargfvaafdaetgklkwrlyvprnknepdhaadnllmka 260
OY 143 --ETWGDYDARMMTG--GAMQITVDPTNLVHGSTAVGPASEHQRPGTILY--- 193
OY 261 ykewgp-kgawrgqg99glwdslllydprdllylgvngsgspwkytriseqgnllfgs 309
OY 194 -GTN-----TRAVRPTGELIWMHQTLPRDNDQECTFEEMAVTVNDVOPSTEMEGLOST 247
OY 310 gisnllfgslvalkpegeywhfgaepmdqdwqysvqimclampvk-----EN 297
OY 248 NPNAATGE-RRVLTGVPCKTGMNQPAETGEELMARDINYONMIESID----- 358
OY 359 -----gemrhlvnap-kngfflyldakqglfsgkpylyenwangldpvtgtrpnydp 411
OY 298 GIYVNEDEALILKELVEYDYCPTEFGSDMPASALNPDGSIYFIPLNVCY----- 348
OY 412 gilyclng-----kfwygl-pgplgahmfaamayspckllylpaqvyrfvydpqkgfkahdswnlglmklgl 463
OY 349 -----DMNAVDEETSMGVYNTSNTKLPQKDMIGRIDALIDISTGRTLSVERAANYSPLSTGGVLENGCT 522
OY 464 kphadswngylmktknglpdpe-artaylkdlyllawdpykmetvklchxgprng 522
OY 405 VSTGGVLENGCTRYFRALSQETGELMOTRLATVASGOALSYEVDMQYVAI---AG-----GGVSYGS 468
OY 523 llatgdlifqglangefnaydatngsdlkyfdaqsqliapmtysvngkyvavewg 582
OY 459 ----IAGGVSYGS 469
OY 583 gilylmgvgvgrts 597

```

RESULT 7
AAR20192 standard; protein: 738 AA.

AA20192;
16-APR-1992 (first entry)

XX DT ADH complex protein (mol.wt. 72.000).
XX DE
XX KM Alcohol dehydrogenase; acetic acid; fermentation.
XX KW Acetobacter altoacetigenes NH-24.
XX OS
XX PN JF03266988-A.

```

Query Match      16.1%; Score 407.5; DB 13; Length 738;
Best Local Similarity 23.7%; Pred. No. 4.7e-27;
Matches 140; Conservative 68; Mismatches 219; Indels 163; Gaps 14;

OY      10  SACAAALLAAPAPFAQVTPYDTE-----LLANPPAGEWLSYGQNCENRHSPLTOYTT 61
      17  tgcctcaaalisygatasdgqatgaallhaddhpnmytgyrtysdqrsplqlnqr 76
OY      62  ENVGQQLVWAGMOPGKVOY-TPLIHGVNVL-----A 94
      77  srvgvnlklawyldldtnrgqgtrplvldgymyattnsmnkavdaatgllwsgyprypg 136
OY      95  NI----- 96
      137  niadggcdctnrgaaywnkgvyfgtfdgrllaladkgtklwsvntllppeeajgkrsy 196
OY      97  -----VANKVIVASTCOYSPFGC-FVSGHDASTGEIMRNFPIRAGEED----- 142
      197  tvdgaqrlakgyrlign-ggsefgarftvsafdaetgkvdcwrtllfvpknepdaasds 254
OY      143  -----ETWGNDYEARMWMTGA-----WQGITDPTVNTLVHGSTAVGPASETOR 185
      255  vlmnkaytws-----ptgawtqggsggtcwasivaypdradivlylgvngspwmykyr 307
OY      186  GTPGCTLYGTNTEFAVRPTGEIYRHRHOTLPBDMWDCETFEKMWTVWVDVQPSHEMGIO 245
      308  segkgdnlftglvalkpetgeyvvnhqtecpmdwdttsdqimcltdlpr----- 357
OY      246  SLPNPAATGERVVLGVGCVCKTGTMMQDAETGEELMARDTNTNMIESLD-ENGIVTVNE 304
      358  -----ngelthvvharkngftfyllidaktgefisgkuyvnyvnsagldpkrtpilynp 410
OY      305  DALILKELVEYDVCPFTLGGRDMPASALINPDSGTFEPLPANNCYIDMAVDOEFT-SMDVY 363
      411  dalyltlltckewyglpbgdlgghnfafamafspkctilyltpqgqpfilytnvgvgfcpbhs 470
OY      364  NTS---NWTKLPG-----KDMGRIDALIDISTGRTLMSVERAANAANSPVLSSTGGV 412
      471  nlgldmktkvglpdspeakgafvkdlygwvavacdpkqgaawrvdhkbpwngslatggdl 530
OY      413  LFNGGTRRYRRAISOETGETLMQTRILAVASQALSYEWDGKQVVAALAG 462
      531  lfnglanelthaydatngsdllhfaadsgliappvtylangkqyvaavevg 580

```

```

RESULT      8
ID          AARI13993 standard; Protein; 738 AA.
XX
AC          AARI13993;
XX
DT          09-DEC-1991 (first entry)
XX
DE          A.altoacetigenes membrane-bound ADH 72KD sub-unit.
XX
KW          alcohol dehydrogenase complex; carboxylic acid production.
XX
OS          Acetobacter altoacetigenes.
XX
PN          EP448969-A.
XX
PD          02-OCT-1991.
XX
PF          26-FEB-1991; 91EP-0102793.
XX
PR          26-MAR-1990; 90JP-0073440.
XX          26-FEB-1990; 90JP-0042391.
PA          (NAKA-) NAKANO VINEGAR KK.
XX
PI          Tamaki T., Takemura H., Tayama K., Fukaya M., Okumura H.;
PI          Kawamura Y.
XX
DR          WPI: 1991-289462/40.
XX          N-PSDB; AAQ13580.
XX
PT          Gene for membrane-bound alcohol dehydrogenase complex - obd.
PT          from Acetobacter altoacetigenes, used for prodn. of enzyme for
PT          converting alcohol to acid
XX
PS          Disclosure; Fig 3; 36pp; English.
XX
CC          Total DNA was prepared from A.altoacetigenes MH-24, digested with
CC          PstI and SmaI and ligated to PstI-SmaI cleaved pUC18. The ligation
CC          mixture was used to transform E.coli JM109. Probes were designed
CC          based on the N-terminal amino acid sequence of the ADH complex
CC          isolated from A. altoacetigenes (see AAQ13582-Q13584). The gene
CC          directly sequenced N-terminal region of the purified 72kd sub-unit
CC          corresponds to the sequence beginning at residue 36 of the deduced
CC          sequence. This suggests that the first 35 N-terminal amino acids
CC          form a leader peptide involved in secretion of the mature 72kd
CC          protein. The deduced amino acid sequence has 77 per cent homology
CC          with the same enzyme from A. acetii K6033. See also AAQ13581.
XX
XX          Sequence 738 AA;
XX
Query Match      16.1%; Score 406.5; DB 12; Length 738;
Best Local Similarity 23.7%; Pied.No.5.8e-27;
Matches 140; Conservative 67; Mismatches 220; Indels 163; Gaps
OY      10 SGCALALLAARFAFQVTPWDE-----LLANPPAGFMISYCQNGENRHSPLQTIT 61
DB      .::.: : : : : : : : : : : : : : : : : : : : : : : : : : :
OY      17 tagtcaalissyvatmasadggagatgaalshadhpnmntlyrgtydsdqysplqlnr 76
DB      ||| | : : : : : : : : : : : : : : : : : : : : : : : :
OY      62 ENVGQLDILVARGKQGCKGVY--TPIIHOCVMYL-----A 94
DB      ||| | : : : : : : : : : : : : : : : : : : : : : : : :
OY      77 snvgnlklawyldldnrgsgelptlvldgvmlyatltnwmkavdaatgklwsysdprryg 136
DB      ||
OY      95 NI-----
DB      137 niaadgcctvmrngaayngkyfgtfqgrllaldaktgylvsvncltppeaelgkqrsy 196
OY      97 -----VANGLIVAGSTCGSPFGC--FVSGHDSATGEELMKNFTIPRAGEEGC----- 142
DB      : : : : : : : : : : : : : : : : : : : : : : : : : :
OY      197 tvdgaprltikgvilgn--ggsetlgargvtasfaetgkvdrftfvpnphnepdaasds 254

```

```

OY 143 -----ETWGNDEARMGCA-----WGQITDPVTNLVHGSSTAVGASFOR 185
XX      |||      |||      |||      |||      |||      |||      |||
DB 255 vlmkayqtws-----pugawtrggggtvdswdsvlyqpvadlvlylgvngspwnykyr 307
OY 186 GTPGGLVSTNRFARPDTEGEIWRHQTLPRDNMOECTFEKAMVNTNVQSTMEGLQ 245
XX      |||      |||      |||      |||      |||      |||      |||
DB 308 segkgnllgslvalkpetgeywhfgeipmdqwdtsdqqlmclidpr-----357
OY 246 SIMPNAATGERRLVGCCKTGMQPDATGEFLMARDTNYONMIESID-ENGIVYNE-304
XX      |||      |||      |||      |||      |||      |||      |||
DB 358 -----ngethvlyharknngfyidantgetisgknyvvnwasgldpktrpiyng 410
OY 305 DALIKELDEVDYCEFTLGGDRMPSALNPDGTYTPIPLANNCYDMMAVDQET-SMDVY 363
XX      |||      |||      |||      |||      |||      |||      |||
DB 411 dalytltygkewyigpdlgghnaamatspkiylyipagqpyllytnqvgfttprdsaw 470
OY 364 NTS---NWTKLPG-----KDMIGRIDIDISTGRTLMSVERAANYSPVLSTGGV 412
XX      |||      |||      |||      |||      |||      |||      |||
DB 471 nlglamkvqipdspeakqafvkdikgylvawdpqkgaearvthkpgwngllatggl 530
OY 413 LEFGSTDRYFRALSOETGETLWQRLAIVASQAISYEVDMQVVAIAGG 462
XX      |||      |||      |||      |||      |||      |||      |||
DB 531 lfgqlangefhayaatngsdlfhaadsgliappvlylangkqyvaevg 580

RESULT 9
AAW95019
ID AAW95019 standard; Protein; 740 AA.
XX
AC AAW95019;
DT 21-MAY-1999 (first entry)
XX
DE Sorbitol dehydrogenase (SDH) protein sequence.
XX
XX Sorbitol dehydrogenase; SDH; open reading frame; ORF2; L-sorbose;
KM D-sorbitol; vitamin C; enzyme.
XX
OS Glucanobacter suboxydans.
XX
FH Key location/Qualifiers
FT Peptide 1..24
FT /note= "signal sequence"
FT Protein 25..740
FT /note= "mature protein"
XX
PN EP897984-A2.
XX
PD 24-FEB-1999.
XX
PF 13-AUG-1998; 98EP-0115231.
XX
PR 21-AUG-1997; 97EP-0114432.
XX
PA (HOFF) HOFFMANN LA ROCHE & CO AG F.
XX
PI Hosino T, Miyazaki T, Ojima S, Shinjoh M, Tomiyama N.
XX
XX WPI; 1999-134646/12.
XX
DR N-PSDB; AAX21501.
XX
PT New D-sorbitol dehydrogenase gene and recombinant protein - useful
PT for production of L-sorbose, an intermediate in vitamin C production
XX
PS Claim 1; Fig 3A-D; 39pp; English.
XX
XX This represents a sorbitol dehydrogenase (SDH) protein. The DNA
XX encoding the SDH enzyme also encodes an open reading frame (ORF2)
XX product upstream of the SDH open reading frame, needed for SDH activity
XX in vivo. Host cells transformed by a vector comprising the SDH DNA
XX sequence are used for the recombinant expression of the sorbitol
XX dehydrogenase. Recombinant D-sorbitol dehydrogenase is useful for
XX producing L-sorbose from D-sorbitol. L-sorbose is an important

```

```

CC Intermediate in vitamin C production.
XX
SQ Sequence 740 AA;
XX
Query Match 10.8%; Score 274.5; DB 20; Length 740;
XX Best Local Similarity 20.6%; Pred. No. 2.8e-15;
XX Matches 143; Conservative 70; Mismatches 167; Indels 313; Gaps 33;
OY 23 AOVTEVTELLANPAGEMTSYSGNDENRHSPLTOITENYGOLOL--VWARGH--QPG 78
XX      |||      |||      |||      |||      |||      |||      |||
DB 86 sgvpamapqgsanpargdvayaygrtdhqtyspslsetipensaskikvafvlyhtgyspfp 145
OY 79 KV-----QVTPPL-IHDGYVYLANIVANGVIVAGSTCOYSPFGCFVSGHDSATGEELMR--130
XX      |||      |||      |||      |||      |||      |||      |||
DB 146 qynkwaeattpikvgyd-lycsam-ndiikl-----dpagkqkwrtn 187
OY 131 -----NYF-----IPRA-----137
XX      |||      |||      |||      |||      |||      |||      |||
DB 188 vdvkynslpylaackgvtlyftsvvpeggqpcmrlliegtldmrliavdaetgdfcpnfigh 247
OY 138 -----GE-----EGDETM-----GNDYEARM 154
XX      |||      |||      |||      |||      |||      |||      |||
DB 248 gggvnlmgjgsevpqfvmptappvngvvvvnhevldgqtrwapsvgrlydaesgkf 307
OY 155 TGAN-----GQITDPVTNLVHGSSTAVGASFORGTPGGLVSTNRF-----198
XX      |||      |||      |||      |||      |||      |||      |||
DB 308 vwadvnnsgrspay-rvclvaverltpjldtrrg-----gsrlprdrnsaady 359
OY 199 -FAVRPD-----TGEIWRHQTLPRDNMOECTFEKAMVNTNVQSTMEGLQ 242
XX      |||      |||      |||      |||      |||      |||      |||
DB 360 ysalsdaenkvsavvaldvktgsprwvfglahkdvwdydgagcl-----ndmp 411
OY 243 GIGSINPNAATGERRLVLCVCKTGMQPDATGEFLMARDTNYONMIESIDEN-----297
XX      |||      |||      |||      |||      |||      |||      |||
DB 412 g-----pdqgtvpallm---pkrgqtfvldrttg-----pdlpveerpasp 452
OY 298 GIV-----TVNEDAI-----LKELD-----312
XX      |||      |||      |||      |||      |||      |||      |||
DB 453 gylpgdprspkpwsgvpallrvpdlketdmwgmsspidqfclfrfranyggetfpssv 512
OY 313 ----VEYDVCPTFGGRVMPALNPDGTYTPIPLN-NVCYDMMAVDQETSM-----360
XX      |||      |||      |||      |||      |||      |||      |||
DB 513 dkpwley---pnyngsdvsgmsypqsgllianwltmptydglvrckkadsjgimpdd 569
OY 361 -----DYNTSNTTKLPPCKDMIGRIDIDISTG-R 390
XX      |||      |||      |||      |||      |||      |||      |||
DB 570 pntkpggggaegngandfygylvtpfwdgy-lgmmcnrpp-----ymtaltadmkgqk 624
OY 391 TLMS-----VERAANYSPVLSTGGVLENG-GTDPYFRALSOETG 430
XX      |||      |||      |||      |||      |||      |||      |||
DB 625 vlgwqhlptarangpwglptglpweilgtpnngsvlvvggllifgaatdnqiraidentg 684
OY 431 ETLWOTRIATVASQAISYEVDMQVVAIAGG 463
XX      |||      |||      |||      |||      |||      |||      |||
DB 685 kvwsavllpgsganpmtyeanghyvalmagg 717

RESULT 10
ABG24430
ID ABG24430 standard; Protein; 443 AA.
XX
AC ABG24430;
DT 18-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #24421.
XX
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX food supplement; medical imaging; diagnostic; genetic disorder.
XX
XX Homo sapiens.
XX
OS

```

PN WO200175067-A2.
 XX 11-OCT-2001.
 XX 30-MAR-2001; 2001WO-US08631.
 XX 31-MAR-2000; 2000US-0540217.
 XX 23-AUG-2000; 2000US-0649167.
 XX (HYSE-) HYSEQ INC.
 XX
 XX Drmanac RT, Liu C, Tang YF;
 XX WPI; 2001-639362/73.
 XX N-PSDB; AAS88617.
 XX
 XX New isolated polynucleotide and encoded polypeptides, useful in
 XX diagnostics, forensics, gene mapping, identification of mutations
 XX responsible for genetic disorders or other traits and to assess
 XX biodiversity -
 PS Claim 20; SEQ ID NO 54789; 103pp; English.
 CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. AAG00010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pcl_sequences.
 XX
 SO Sequence 443 AA;

Query Match 5.4%; Score 137; DB 22; Length 443;
 Best Local Similarity 30.0%; Pred. No. 0.0019;
 Matches 42; Conservative 18; Mismatches 38; Indels 42; Gaps 7;

QY 6 LMAS-----AGALALAPAFAYPTVDELANPPAGEWISGONENYRHSPLNQ 58
 DB 168 lwagfndpgeingflsdairpaee-ispvadt-----dwpaygrgqgfrfslpkq 218
 QY 59 ITTEWVGLOLVMA-----RGMQPKV--QVTPLIHDSVMTLANIVANGVIYAGSICQ 109
 DB 219 inedvthnikawwfrtgdkvqpnidpgetnevtprtkvgdtlylc-----tah 266
 QY 110 YSPGCEVSGHDSATGEBELN 129
 DB 267 qrlfal-----daasgfkew 281

RESULT 11
 ID AAG81738 standard; Protein; 715 AA.
 AC AAG81738;
 XX
 XX 03-SEP-2001 (first entry)
 XX

DE S. epidermidis open reading frame protein sequence SEQ ID NO:570.
 XX Staphylococcus epidermidis SRI strain; infection; diagnosis;
 XX Staphylococcus epidermidis.
 XX Staphylococcus epidermidis.
 XX WO200134809-A2.
 XX 17-MAY-2001.
 XX 09-NOV-2000; 2000WO-US30782.
 XX 09-NOV-1999; 99US-0164258.
 XX (GLAX) GLAXO GROUP LTD.
 XX Kimmmerly WJ;
 XX WPI; 2001-316495/33.
 XX N-PSDB; AAH52588.
 XX
 XX Nucleic acids encoding polypeptides from Staphylococcus epidermidis,
 XX useful for vaccinating against infections, e.g. endocarditis -
 PS Claim 18; Page 189; 2188pp; English.

CC AAH5304 to AAH53970 represent nucleic acids (I) encoding polypeptides
 CC (II), given in AAG81454 to AAG83120, from Staphylococcus epidermidis.
 CC (I) and (II) can have antibacterial activity and therefore can be used
 CC in vaccination. The nucleic acids (I) may be used to produce the
 CC S. epidermidis polypeptides (II) via the production of vectors
 CC containing them which are used to produce hosts cells which express the
 CC polypeptides. The polypeptides (II) (and/or nucleic acids) may then be
 CC used to vaccinate subjects and to raise antibodies against the bacteria.
 CC The polypeptides may also be used to assay for other inhibitors of their
 CC activity and therefore identify compounds that may be used for the
 CC treatment of S. epidermidis infections, e.g. endocarditis. AAH53971 to
 CC AAH55098 represent specifically claimed S. epidermidis genomic DNA
 CC polynucleotide sequences from the present invention. AAH55091 to
 CC N.B. The present invention specifically claims all the polynucleotide
 CC sequences given in the sequence listing of the present specification,
 CC however the sequence listing only goes up to SEQ ID NO:4454 so even
 CC no sequences are present for SEQ ID NO:4455 to 4472.
 XX
 SO Sequence 715 AA;

Query Match 4.8%; Score 122.5; DB 22; Length 715;
 Best Local Similarity 23.1%; Pred. No. 0.0731;
 Matches 99; Conservative 43; Mismatches 155; Indels 131; Gaps 24;

QY 26 TPVTDLANPPAGEWISGONENYRHSPLTQIT-----TENVGLOLVMAWGMO- 76
 DB 309 tptt-----knpitgkxgkgeptkvtkqpydelvhygseikpghkdefdpnapysqe 364
 QY 77 --PGR-----VOVPTLIHDSVMTLANIVANGVIYAGSICQYSPGCEVSGHDSATG 125
 DB 365 dvpgkpygknpdgtgevtvtpvdd-----vdkygp-----vdgdpitst 402
 QY 126 BEL--WRNYFIP-----RAGEGDETWNDEYFARMWTG-----AMQGITYDP 165
 DB 403 eelpfdkkrefdpnaplpgtekavvgkgeptkltitptknpitgkxgkgeptekvtkq 462
 QY 166 VTNLVHYGSTAVGP-----ASETQRCIPGGLTGTNTRFAVRDPTGEIWRH 212
 DB 463 vdelvhygseikpghkdefdpnapkysgtltqpgkpy-----vknptdtegvv--- 509
 QY 213 QTLPRDNWDOCTEFEMWVTN--VDVQPTSTEMEGL--OSINPMNATGERR-LTGV 263

Db 510 -lppvdd-----vtkygvddgplsteeipfdkkrfepnllkpgseervkqgpe 558
 QY 264 -CKT-GTMMQFAEENGELMARDNYCNMIESIDENCIYVNDALIKELDYEVCTFE 321
 Db 559 gktlttpttknplqekvggeprekltkqvde--lteygggeekpkhkdelfdpnkp 616
 QY 322 LGGSDWPS--AALNPDGS-IVFIPLNANCY-----DMAVADGF--FTSMYVNTSNVTK 370
 Db 617 gsgedvpgkqgvknptdgtvvpvddvtkygvddgplsteeipdkkrfepnd---- 672
 QY 371 LPPGKDMI 378
 Db 673 lpgkerv 680

RESULT 12
 ABG25909
 ID ABG25909 standard; Protein; 948 AA.
 AC ABG25909;
 DT 18-FEB-2002 (first entry)
 DE Novel human diagnostic protein #25900.
 KW Human: chromosome mapping; gene mapping; gene therapy; forensic;
 food supplement; medical imaging; diagnostic; genetic disorder.
 OS Homo sapiens.
 PN WO200175067-A2.
 PD 11-OCT-2001.
 PF 30-MAR-2001; 2001WO-US08631.
 PR 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 PA (HYSE-) HYSEQ INC.
 PI Drmanac RT, Liu C, Tang YT;
 DR N-PSDB; AAS90096.
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity
 PS Claim 20; SEQ ID No 56268; 103bp; English.

The invention relates to isolated polynucleotide (I) and
 polypeptide (II) sequences. (I) is useful as hybridisation probes,
 polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 and gene mapping, and in recombinant production of (II). The
 polynucleotides are also used in diagnostics as expressed sequence tags
 for identifying expressed genes. (I) is useful in gene therapy techniques
 to restore normal activity of (II) or to treat disease states involving
 (II). (II) is useful for generating antibodies against it, detecting or
 quantitating a polypeptide in tissue, as molecular weight markers and as
 a food supplement. (II) and its binding partners are useful for treating
 disorders involving aberrant protein expression or biological activity.
 The polypeptide and polynucleotide sequences have applications in
 diagnostics, forensics, gene mapping, identification of mutations
 responsible for genetic disorders or other traits to assess biodiversity
 and to produce other types of data and products dependent on DNA and
 amino acid sequences. ABG00010-ABG30377 represent novel human
 diagnostic amino acid sequences of the invention.
 Note: The sequence data for this patent did not appear in the printed
 specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published_pcl_sequences.
 XX Sequence 948 AA;
 SQ
 Query Match 4.8%; Score 122.5; DB 22; Length 948;
 Best Local Similarity 34.4%; Pred. No. 0.11; Indels 25; Gaps 5;
 Matches 33; Conservative 14; Mismatches
 QY 6 ILWAS-----AGLALLAPAFQVTPVDELIANPAGEMISYGCONEYRHSPLTO 58
 Db 310 ltwaqndpgeingtltsadcpaea-ispvdaq-----dwpaygrnqegqritspla 360
 QY 59 ITTENYGOLOLWVA-----RCMQPGKV--QVTPPL 85
 Db 361 lndnvhmlkeavfrtgvtkgndpgseitnevpl 396

RESULT 13
 ABG21573
 ID ABG21573 standard; Protein; 1510 AA.
 AC ABG21573;
 DT 18-FEB-2002 (first entry)
 DE Novel human diagnostic protein #21564.
 KW Human: chromosome mapping; gene mapping; gene therapy; forensic;
 food supplement; medical imaging; diagnostic; genetic disorder.
 OS Homo sapiens.
 PN WO200175067-A2.
 PD 11-OCT-2001.
 PF 30-MAR-2001; 2001WO-US08631.
 PR 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 PA (HYSE-) HYSEQ INC.
 PI Drmanac RT, Liu C, Tang YT;
 DR N-PSDB; AAS85760.
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity
 PS Claim 20; SEQ ID No 51932; 103bp; English.

The invention relates to isolated polynucleotide (I) and
 polypeptide (II) sequences. (I) is useful as hybridisation probes,
 polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 and gene mapping, and in recombinant production of (II). The
 polynucleotides are also used in diagnostics as expressed sequence tags
 for identifying expressed genes. (I) is useful in gene therapy techniques
 to restore normal activity of (II) or to treat disease states involving
 (II). (II) is useful for generating antibodies against it, detecting or
 quantitating a polypeptide in tissue, as molecular weight markers and as
 a food supplement. (II) and its binding partners are useful for treating
 disorders involving aberrant protein expression or biological activity.
 The polypeptide and polynucleotide sequences have applications in
 diagnostics, forensics, gene mapping, identification of mutations
 responsible for genetic disorders or other traits to assess biodiversity
 and to produce other types of data and products dependent on DNA and
 amino acid sequences. ABG00010-ABG30377 represent novel human

CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 1510 AA;
 SQ

Query Match
 Best Local Similarity 4.8%; Score 122.5; DB 22; Length 1510;
 Matches 33; Conservative 14; Mismatches 24; Indels 25; Gaps 5;

QY 6 LMAS-----AGALALAAFAFAQVPTVDLLANPAGEWISYCONENRHSPLNO 58
 Db 310 ltwagfndpgeingtltsadapaea-ispvadq-----dwpaygrngqgrfslpkq 360
 QY 59 ITTENGGOLQLYMA-----RGMOGRKV--QVTPPL 85
 Db 361 lmadvnhlkeawfrtgdkvqgndpgeitnevtpi 396

RESULT 14

ID AAG83007 standard; Protein; 910 AA.
 AC AAG83007;

DT 03-SEP-2001 (first entry)

DE S. epidermidis open reading frame protein sequence SEQ ID NO:3108.
 KW Staphylococcus epidermidis SRI strain; infection; diagnosis;
 XX vaccination; endocarditis.

OS Staphylococcus epidermidis.

PN MO200134809-A2.

PD 17-MAY-2001.

PF 09-NOV-2000; 2000WO-US30782.

PR 09-NOV-1999; 99US-0164258.

PA (GLAX) GLAXO GROUP LTD.

PI Kimerly MJ;

DR WPI: 2001-316495/33.

DR N-PSDB; AAH53857.

XX Nucleic acids encoding polypeptides from Staphylococcus epidermidis,
 XX useful for vaccinating against infections, e.g. endocarditis -
 XX Claim 18; Page 819; 2188pp; English.

CC AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides
 CC (II), given in AAG81454 to AAG83120, from Staphylococcus epidermidis
 CC in vaccination. The nucleic acids (I) may be used to produce the
 CC S. epidermidis polypeptides (II) via the production of vectors
 CC containing them which are used to produce hosts cells which express the
 CC polypeptides. The polypeptides (II) (and/or nucleic acids) may then be
 CC used to vaccinate subjects and to raise antibodies against the bacteria.
 CC The polypeptides may also be used to assay for other inhibitors of their
 CC activity and therefore identify compounds that may be used for the
 CC treatment of S. epidermidis infections, e.g. endocarditis. AAH53971 to
 CC AAH5090 represent specifically claimed S. epidermidis genomic DNA
 CC AAH5098 represent oligonucleotide sequences and primers which are used
 CC in the exemplification of the present invention. AAH5091 to
 CC N.B. The present invention specifically listing all the polynucleotide
 CC sequences given in the sequence listing of the present specification,

CC however the sequence listing only goes up to SEQ ID NO:4454 so even
 CC though sequences are given in the disclosure for SEQ ID NO:4465 to 4472,
 CC no sequences are present for SEQ ID NO:4455 to 4464.
 XX
 SQ Sequence 910 AA;

Query Match
 Best Local Similarity 4.8%; Score 122; DB 22; Length 910;
 Matches 92; Conservative 37; Mismatches 146; Indels 98; Gaps 21;

QY 26 TPYTDLLANPAGEWISYCONENRHSPLTOIT--ENV-----GOLQLYMARQMO- 76
 Db 231 tptt-----knpitgekvgepelekvkqpvdeitelyggeelkpgkhdefdpnarkgsge 286
 QY 77 --PGR-----VOYTPLIHOGVMYLANIYANGIYVAGSTCOYSPGCFVSGHDSATG 125
 Db 287 dvpqkpvknpdgtgevtlprvdd-----vdkygp-----vdgdpitst 324
 QY 126 EEL--WRNYFIP-----RAGESEDETWGNDEYARWMTG-----AMGQITYP 165
 Db 325 eelpfdkkrefdpnlapgtkvvgkqepgtklitpttknpltgkvgepelekvkq 384
 QY 166 VTNLVHRTGAVGPA--SETORATPGTLYGNTNFAVR--PDGTETIYWRHQTLPNDMDQ 222
 Db 385 vdelvhyggeelkpgkhdefdpnarkgsgevdvqkpgknpdgtgevv-----ltpvdd--- 437
 QY 223 ECTFEMWVTV--VDYOPSTEMEG--OSINPNATG-ERRVLTGVP-CRT-GTMMQ 271
 Db 438 -----vdkypvqgdplsteeelpfdkkrefdpnlapgtkvvgkqepgtklitptt 490
 QY 272 FDAETGEPIFMAADTYQNMFTSIDENGIYVNEDEALIKELDYDVCPTFLGGRWPS-- 329
 Db 491 knplgkvgepelekvkqpvde--lvhyggeelkpgkhdefdpnarkgsgevdvqkpg 548
 QY 330 AALNPDSGIYFIP 342
 Db 549 gvknpdgtgevtlpr 561

RESULT 15

ID AAG82914 standard; Protein; 696 AA.

AC AAG82914;

DT 03-SEP-2001 (first entry)

DE S. epidermidis open reading frame protein sequence SEQ ID NO:2922.
 KW Staphylococcus epidermidis SRI strain; infection; diagnosis;
 XX vaccination; endocarditis.

OS Staphylococcus epidermidis.

PN MO200134809-A2.

PD 17-MAY-2001.

PF 09-NOV-2000; 2000WO-US30782.

PR 09-NOV-1999; 99US-0164258.

PA (GLAX) GLAXO GROUP LTD.

PI Kimerly MJ;

DR WPI: 2001-316495/33.

DR N-PSDB; AAH53764.

XX Nucleic acids encoding polypeptides from Staphylococcus epidermidis,
 XX useful for vaccinating against infections, e.g. endocarditis -
 XX

PS Claim 18; Page 763-764; 2188bp; English.

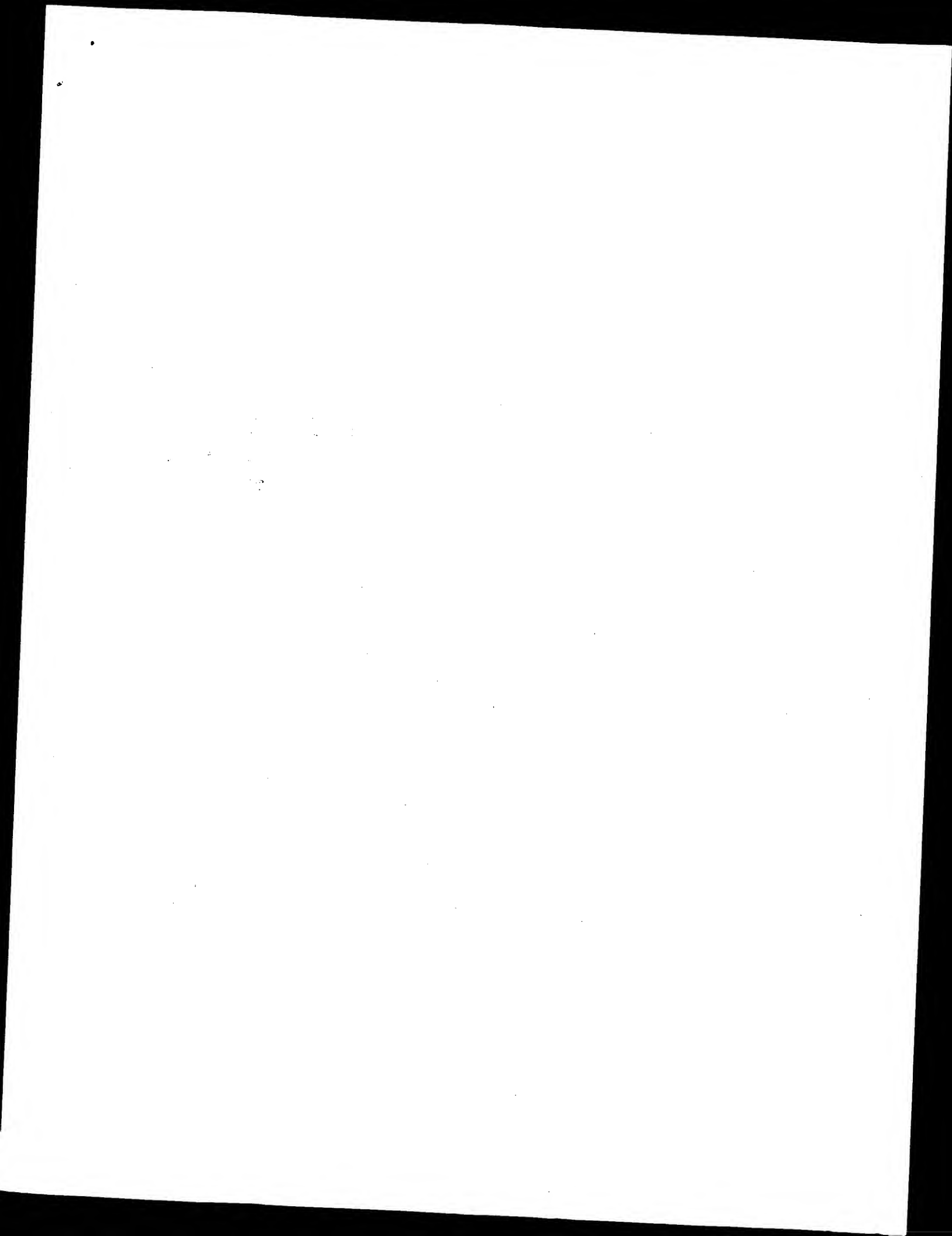
XX AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides
CC (II), given in AAH53970, from *Staphylococcus epidermidis*.
CC (II) and (II) can have antibacterial activity and therefore can be used
CC in vaccination. The nucleic acids (I) may be used to produce the
CC S. epidermidis polypeptides (II) via the production of vectors
CC containing them which are used to produce host cells which express the
CC polypeptides. The polypeptides (II) (and/or nucleic acids) may then be
CC used to vaccinate subjects and to raise antibodies against the bacteria.
CC The polypeptides may also be used to assay for other inhibitors of their
CC activity and therefore identify compounds that may be used for the
CC treatment of S. epidermidis infections, e.g. endocarditis. AAH53971 to
CC AAH55090 represent specifically claimed S. epidermidis genomic DNA
CC polynucleotide sequences from the present invention. AAH55091 to
CC AAH55098 represent oligonucleotide sequences and primers which are used
CC in the exemplification of the present invention.
CC N.B. The present invention specifically claims all the polynucleotide
CC sequences given in the sequence listing of the present specification,
CC however the sequence listing only goes up to SEQ ID NO:4454 so even
CC though sequences are present for SEQ ID NO:4455 to 4464.
XX no sequences are present for SEQ ID NO:4455 to 4464.

SO Sequence 696 AA;

Query Match 4.8%; Score 121.5; DB 22; Length 696;
Best Local Similarity 23.5%; Pred. No. 0.087;
Matches 101; Conservative 48; Mismatches 147; Indels 133; Gaps 25;

OY 26 TPVTDELLANPPAGWISYQONQENYRHSPLQITTT---ENV-----GQLQLYWARGMO- 76
DB 103 tptt---knplgkvggeptektikpvdeltgkgeikpghkdefdnapkysgse 158
OY 77 --PGK-----VOYTPLIHDGVMITANIYANGVIVAGSTCOYSPFGCFVSGHDSATG 125
DB 159 dvpgkpgyknptdgtvtpvdd-----vtkygp-----vdgdpiltst 196
OY 126 EEL---KKNYFIP-----RAGEGDETWGNDYEAKMTG-----AMGQITYDP 165
DB 197 eelipdkkretlnpnapgtkvgkgepgtkltitpknplgkvggeptektikp 256
OY 166 VTNLIVHYGSTAVGPA--SETORGTGPGTILGTNTREAVR--PDGELIWRHQTLPDNDMDQ 222
DB 257 vdeivhyggeelkpgkdeidfnapkgsgeadvpgkpgvknptdgtv-----tpvdd--- 309
OY 223 ECTFEEMWNTN---VDVOSTEMEGT-----OSINPNAATG--ERRVLTGVP--CKT--GTMMQ 271
DB 310 -----vtkygpgvdgdsitsteelipdkkretlnpnapgtkvgkgepgtkltitp 362
OY 272 FDATEGELIWARDNTYNQNMIESIDENGIVTNEADAILKELDEYDVCPTFLGGRDWPSSAA 331
DB 363 knplgkvggeptektikpvd--lvhyggeidpghkdefd-----pnsp 409
OY 332 L-----NPDSC-IYFIPLINVCYDMAAVDOEFTSMQ-----VYNTSNVT 369
DB 410 vdsktevpgkpgvknptdgtvtpvddvtkygpgvknplstteelipdkkrtvfnpd--- 466
OY 370 KLPFGKDMI 378
DB 467 -lkpggeerv 474

Search completed: May 24, 2002, 10:19:16
Job time: 379 sec



walick-934-95.pep.ra1

Fri May 24 11:27:35 2002

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model
May 24, 2002, 10:14:47 : Search time 39.53 Seconds
(without alignments)
291.649 Million cell updates/sec

Title: WALICK-934-95.PEP
Sequence: 1 MKPTSLMASGALALALAAP.....GMOYVIALGGVSYGSLNS 472

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5
Searched: 231628 seqs, 24425594 residues
Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: Issued Patents AA:*
2: /cgn2_6/prodata/2/1aa/5A.COMB.pep:*
3: /cgn2_6/prodata/2/1aa/5B.COMB.pep:*
4: /cgn2_6/prodata/2/1aa/6A.COMB.pep:*
5: /cgn2_6/prodata/2/1aa/6B.COMB.pep:*
6: /cgn2_6/prodata/2/1aa/PCTUS.COMB.pep:*
7: /cgn2_6/prodata/2/1aa/Backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	413	16.3	754	4 US-09-296-284-4	Sequence 4, Appli
2	407.5	16.1	738	1 US-07-985-458-3	Sequence 3, Appli
3	404	16.0	720	4 US-09-296-284-25	Sequence 25, Appli
4	279.5	11.0	739	3 US-09-136-251-2	Sequence 2, Appli
5	105	4.2	1012	1 US-08-219-262B-10	Sequence 10, Appli
6	105	4.2	1012	3 US-09-031-655-10	Sequence 2, Appli
7	103.5	4.1	500	2 US-08-987-519-2	Sequence 6, Appli
8	100	4.0	532	3 US-08-737-336-6	Sequence 1, Appli
9	97.5	3.9	500	2 US-08-987-519-1	Sequence 2, Appli
10	95.5	3.8	548	5 PCT-US93-10541-2	Sequence 19, Appli
11	95	3.8	1012	1 US-08-216-276A-19	Sequence 2, Appli
12	94.5	3.7	512	4 US-09-356-818A-11	Sequence 11, Appli
13	94	3.7	1042	3 US-08-928-361B-15	Sequence 5, Appli
14	94	3.7	1837	3 US-08-928-361B-11	Sequence 2, Appli
15	94	3.7	5215	4 US-09-105-537-2	Sequence 2, Appli
16	93.5	3.7	548	1 US-08-247-902A-2	Sequence 2, Appli
17	93.5	3.7	583	4 US-09-401-476-2	Sequence 4, Appli
18	93	3.7	380	3 US-08-971-782-4	Sequence 4, Appli
19	93	3.7	380	4 US-09-809-026-4	Sequence 2, Appli
20	93	3.7	459	3 US-08-971-782-2	Sequence 2, Appli
21	93	3.7	459	4 US-09-309-026-2	Sequence 15, Appli
22	92.5	3.7	1222	2 US-08-682-517-15	Sequence 9, Appli
23	92.5	3.7	1252	2 US-08-682-517-9	Sequence 3, Appli
24	92	3.6	606	2 US-08-883-534-3	Sequence 3, Appli
25	92	3.6	606	3 US-09-204-764-3	Sequence 4, Appli
26	92	3.6	812	1 US-08-446-794A-4	Sequence 4, Appli
27	91.5	3.6	816	1 US-07-731-157A-4	Sequence 4, Appli

28	91.5	3.6	816	1 US-08-229-444B-2	Sequence 2, Appli
29	91.5	3.6	816	2 US-08-541-780-4	Sequence 4, Appli
30	91.5	3.6	1612	1 US-08-169-927-2	Sequence 2, Appli
31	91	3.6	551	2 US-08-793-229-32	Sequence 32, Appli
32	91	3.6	551	3 US-08-955-957-32	Sequence 29, Appli
33	90.5	3.6	322	1 US-08-216-276A-29	Sequence 2, Appli
34	90.5	3.6	687	5 PCT-US91-09784-2	Sequence 5, Appli
35	90.5	3.6	1721	3 US-08-700-651-5	Sequence 17, Appli
36	90	3.6	484	1 US-08-216-276A-17	Sequence 3, Appli
37	90	3.6	714	2 US-08-472-534-3	Sequence 12, Appli
38	90	3.6	1012	1 US-08-219-262B-12	Sequence 12, Appli
39	90	3.6	1012	3 US-09-031-655-12	Sequence 2, Appli
40	90	3.6	1012	3 US-09-031-655-12	Sequence 2, Appli
41	89.5	3.5	820	1 US-08-291-886-2	Sequence 2, Appli
42	89.5	3.5	820	2 US-08-485-278-2	Sequence 3, Appli
43	89.5	3.5	723	1 US-08-290-937B-3	Sequence 3, Appli
44	89	3.5	723	2 US-08-290-937B-3	Sequence 3, Appli
45	88	3.5	1012	1 US-07-944-943-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-09-296-284-4
Sequence 4, Application US/09296284A
Patent No. 6204040
GENERAL INFORMATION:
APPLICANT: Choi, Eun-Sung
APPLICANT: Lee, Eun-Hae
APPLICANT: Rhee, Sang-Ki
TITLE OF INVENTION: Glucosyltransferase, Sorbitol Dehydrogenase, Genes
FILE REFERENCE: 1533, 0870000
CURRENT FILING DATE: 1999-04-22
NUMBER OF SEQ ID NOS: 87
SOFTWARE: Patent Ver. 2.0
SEQ ID NO 4
LENGTH: 754
TYPE: PRT
ORGANISM: Glucosyltransferase suboxydase
US-09-296-284-4

Query Match	16.38	Score 413	DB 4	Length 754
Best local similarity	23.84	Pred. No. 1.4e-31		
Matches 143	Conservative	72	Mismatches 222	Indels 164
				Gaps 15
OY	14 LALLAAPAFQVTPY	1	1	1
DB	17 LGGAAALAFCAISPYALAEDEGTATNADDPGDMSYGRITSEORXSPLODITKDNASN	76	76	76
OY	67 LQLVWAGMOPKQVY-PTLHNDGVNYL	1	1	1
DB	77 LKLMWHDLDITNGDGTPLVDGMYATNTMSKMLADATKSLMSIDPKYPCNIDR	136	136	136
OY	100 G-----	1	1	1
DB	137 GCGDVTNGAAYNMKVFSTFGRLIALDAKTGLVSYTVPRKAQLOHRSYTVDA	196	196	196
OY	101 -----VYAGSTCOVSPGCFVSGHDSNTGEIMRNFTIPRAGEED	142	142	142
DB	197 PRIAKKVTITNGAERFARG-PTAYDAETGKMDRFTVNPKNRKAASDVMASK	255	255	255
OY	143 --ETWGNDEARWMTG--AMGQITVDPVNLVHVSSTAVGPASETORGTGGLTYGNT	197	197	197
DB	256 APTWKGKMGKQCGGCTWDSLTIDPYDVLVGLVGNSSPNVKKFREGGNLFLGS	315	315	315
OY	198 RFAPVPTGELVWRHDTLPDNDDECTFEAMVTVVQSTEMEGLOSINPAATGE-R	256	256	256
DB	316 IVALNPDTGKTVHPEETPMDDQDYTSVOQITADLDMV-----NGEKR	358	358	358

QY 257 RVLTGVPCKTGTMMQDAETGEFLMARDTNONTMIESIDE-NGIYVNEAIIKELDEX 315
 Db 359 HVLVHAP-KNGFEIITDAKTKFKISGRPTTYENMANGLDPTGPNPNPALTLNGKPM 417
 QY 316 DVCPEPLGGDRMPSALINDPSGIFPLNNVCY-----DMAVDEPFMSMV 362
 Db 418 YGIGDGLGNNFAMAMVSPOTKLYVTPAQOVPFYDDQKGFRAHDSMNLGLDMNKIGL 477
 QY 363 YKNSN---VTKLPKGMIGRIDIDISTRTLMVSERRAANSPVLTSGGGLFNGGT 418
 Db 478 LDNDQOHKADKAQFLKDLKGIYAWDPOKQAFFYDHKGPNNGSLATAGGLFDSLA 537
 QY 419 DRYFALSOETGETLMQRLATVAGSAISEYVDGMOYAL---AG-----GVSYS 468
 Db 538 NGEFHAYDATGKDLFTTPAOSALIPVYTYANGKQYAVAEVNGGLYPPGLGVARIS 597
 QY 469 G 469
 Db 598 G 598

RESULT 2

US-07-985-458-3
 Sequence 3, Application US/07985458
 Patent No. 5344777
 GENERAL INFORMATION:

APPLICANT: Tamaki, Toshimi;
 APPLICANT: Takemura, Hiroshi;
 APPLICANT: Takemura, Kenji;
 APPLICANT: Fukaya, Masahiro;
 APPLICANT: Okumura, Hajime and
 APPLICANT: Kawamura, Yoshiya
 TITLE OF INVENTION: Structural Gene of Membrane-Bound
 TITLE OF INVENTION: Alcohol Dehydrogenase Complex, Plasmid
 TITLE OF INVENTION: Containing the Same And Transformed Acetic Acid
 NUMBER OF SEQUENCES: 4
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Frishaut, Holtz, Goodman & Woodward, P.C.
 STREET: 600 Third Avenue
 CITY: New York
 STATE: New York
 COUNTRY: USA
 ZIP: 10016-2088

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 inch, 0.72 mb
 COMPUTER: IBM PC compatible (NEC PC-9801 ES)
 OPERATING SYSTEM: MS DOS
 SOFTWARE: ASCII FORM

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/985,458
 FILING DATE: 19921203
 CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/658,221
 FILING DATE: 20-FEB-1991
 APPLICATION NUMBER: 73440/1990

ATTORNEY/AGENT INFORMATION:

NAME: Goodman, Herbert
 REGISTRATION NUMBER: 17081
 REFERENCE/DOCKET NUMBER: 910134/HG

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212)972-1400
 TELEFAX: (212)370-1622

TELEX: 236268

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:
 LENGTH: 738 amino acids
 TYPE: amino acid
 TOPOLOGY: unknown

MOLECULE TYPE: protein

NAME/KEY: MATURE PEPTIDE
 LOCATION: 36 to 738
 IDENTIFICATION METHOD: N-terminal sequences of the
 IDENTIFICATION METHOD: purified protein having a molecular weight of about
 ORIGINAL SOURCE: 72,000
 ORGANISM: Acetobacter altoacetigenes
 STRAIN: MH-24
 PUBLICATION INFORMATION:
 AUTHORS: Tamaki, Toshimi;
 AUTHORS: Takemura, Hiroshi;
 AUTHORS: Takemura, Kenji;
 AUTHORS: Okumura, Hajime;
 AUTHORS: Kawamura, Yoshiya;
 AUTHORS: Mshiyama, Makoto;
 AUTHORS: Horinouchi, Sueharu and
 Bepu, Teruhiko
 TITLE: Cloning and sequencing of the Gene Cluster
 TITLE: Encoding two subunits of Membrane-Bound
 TITLE: Alcohol Dehydrogenase from Acetobacter
 JOURNAL: Biochimica et Biophysica Acta.
 VOLUME: 1088
 PAGES: 292-300
 DATE: 1991
 US-07-985-458-3

Query Match

Best Local Similarity 16.1%; Score 407.5; DB 1; Length 738;
 Matches 140; Conservative 68; Mismatches 219; Indels 163; Gaps 14;

QY 10 SAGALLAARAFQVTPVDE-----LLANPACEMWSTYCONENTRHSPLOITT 61
 Db 17 TAGTICALLISGATWASADDDGATGATGALIHADHCHNMWYGRYSDDQSPILQINR 76
 QY 62 ENVOLOLVWARGMOPKQV--TPLIHGQVYL-----A 94
 Db 77 SNGNKLAWYLDLDTRNGQESTPLVDGVMATTNMSMKAVADATGKLSYDPRVG 136
 QY 95 NI----- 96
 Db 137 NIADKCCDTVRGAYNGKVIYGTGRLIALDAKTKLVWSVNTIPPAELGKORSY 196
 QY 97 -----VANGIYVAGSTCOYSPFGC--FVSGHDSATGELMRNFTIPRAGECD----- 142
 Db 197 TVDAPRIAKGRVITGN--GGSEFGARGFVSADFETSKVDMREFVTPNPKNEPDAADS 254
 QY 143 -----ETWGNDEYARWMTGA-----WQIYDPVTLNVHGSYAVGPASETOR 185
 Db 255 VLMKAVQVWS-----PTGAMTQGGGGVWDSIYDPAVDLYLVGNGSPNNYKTR 307
 QY 186 GTPGTLVGTNTPFAVRPTGELVRRHOTLPRDWMDCECFEMWYNNVVOYSTMEGLQ 245
 Db 308 SEGKGNLFLGSLVALKPEGTGYVHFQETPMQMDFTSDQINTIDLP----- 357
 QY 246 SIMPNATGERVLTGVPCKTGTMMQDAETGEFLMARDTNONTMIESIDE-NGIYVNE 304
 Db 358 -----NETHTVIVHARKNGFFIIDAATGTFISGRKYVYVNNASGLDKPCRPIYNP 410
 QY 305 DALIELDVEYDVCPTFGGDRMPSALINDPSGIFPLNNVCYDMAVDEPFMSMV 362
 Db 411 DALYTLTGKEMYGIPDLGHNFAAAAFSPKGLVYIPAQOVPFYDDQKGFRAHDSMNLGLDMNKIGL 477
 QY 364 NTS---NWTKLPKGMIGRIDIDISTRTLMVSERRAANSPVLTSGGGLFNGGT 418
 Db 471 NLGIDMKVGIPIPSPEAKQAFKDLKGIYAWDPOKQAFFYDHKGPNNGSLATAGGLFDSLA 537
 QY 413 LFGNGTDYFRLSOTGETLMQRLATVAGSAISEYVDGMOYAL---AG-----GVSYS 468
 Db 531 LFOGLANGEFHAYDATNGSDLFHRAADSGIIPVYTYANGKQYAVAEVNGGLYPPGLGVARIS 597

Fri May 24 11:27:35 2002

walick-934-95.pcp.ra1

Page 4

Db 685 KVMASVLPGGGQANPMTYANGHQAVALAGG 716

RESULT 5

US-08-219-262B-10

Sequence 10, Application US/08219262B

Patent No. 5788970

GENERAL INFORMATION:

APPLICANT: VAKHARIA, VIKRAM

APPLICANT: SNYDER, DAVID B

TITLE OF INVENTION: CHIMERIC INFECTIOUS BURSAL DISEASE VIRUS

TITLE OF INVENTION: CDNA CLONES, EXPRESSION PRODUCTS AND VACCINES BASED

NUMBER OF SEQUENCES: 15

CORRESPONDENCE ADDRESSES:

ADDRESS: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT

STREET: 1755 S. JEFFERSON DAVID HIGHWAY, FOURTH FLOOR

CITY: ARLINGTON

STATE: VIRGINIA

COUNTRY: USA

ZIP: 22202

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/219,262B

FILING DATE: 29-MAR-1994

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: OBLON, NORMAN F

REGISTRATION NUMBER: 24,618

REFERENCE/DOCKET NUMBER: 2747-047-27

TELEPHONE: (703) 413-3000

TELEFAX: (703) 413-2220

TELEX: 248855 OPAT UR

INFORMATION FOR SEQ ID NO: 10:

SEQUENCE CHARACTERISTICS:

LENGTH: 1012 amino acids

TYPE: amino acid

STRANDEDNESS:

MOLECULE TYPE: protein

ORIGINAL SOURCE:

ORGANISM: Infectious bursal disease virus

STRAIN: OH

US-08-219-262B-10

Query Match

Best Local Similarity 4.2%; Score 105; DB 1; Length 1012;

Matches 84; Conservative 45; Mismatches 130; Indels 120; Gaps 19;

Db 166 VTNLVHGSYAV-----GPAS-----ETORGPGGTLGTWTF 199

Db 1 MTNLMHTQOIVPFRSLMPTTGPASIPDDTLKHTLRSETSTYNLTGVDGSLGVTF 60

Db 200 AVRPDGTGEIWRHQTLPD---NMDOECTFEKMTV-----NVDVOPSTE 240

Db 61 PGFP--GSVYGAHYTLQNSGSGFDQ---MLTLAONLPVSYNYCRLVRSRLTVRSSTL 113

Db 241 MEGIOSINP--NAAGGERVLTGVPCKTGTWMPDAETGEFLVARDTNQMTEST----- 294

Db 114 PGSVYALNCTINAVT-----FQGSLSL-----TDYTINGLMSATANIN 152

Db 295 DENGITVNEADILKELDEYDVOCPTFLGDMPSALNDGSIYFIPLNVC----- 347

Db 153 DKIGNLVGEGYVLSLTSTLSDYVRIGD--PIPAAGLDP-----KLMAICDSSDRP 204

Db 348 -YDMANVQD-EE-----TSMVYNTSNVTKLPKPKDMI-GRIDAIDISTGRTLM- 393

Query Match

Best Local Similarity 4.2%; Score 105; DB 3; Length 1012;

Matches 84; Conservative 45; Mismatches 130; Indels 120; Gaps 19;

Db 166 VTNLVHGSYAV-----GPAS-----ETORGPGGTLGTWTF 199

Db 1 MTNLMHTQOIVPFRSLMPTTGPASIPDDTLKHTLRSETSTYNLTGVDGSLGVTF 60

Db 200 AVRPDGTGEIWRHQTLPD---NMDOECTFEKMTV-----NVDVOPSTE 240

Db 61 PGFP--GSVYGAHYTLQNSGSGFDQ---MLTLAONLPVSYNYCRLVRSRLTVRSSTL 113

Db 241 MEGIOSINP--NAAGGERVLTGVPCKTGTWMPDAETGEFLVARDTNQMTEST----- 294

Db 114 PGSVYALNCTINAVT-----FQGSLSL-----TDYTINGLMSATANIN 152

Db 295 DENGITVNEADILKELDEYDVOCPTFLGDMPSALNDGSIYFIPLNVC----- 347

Db 153 DKIGNLVGEGYVLSLTSTLSDYVRIGD--PIPAAGLDP-----KLMAICDSSDRP 204

Db 205 VYVTADEYOFSSOLIPSGVKTTLFTANIDALSLSVGELLFSQVTHSLEVDVTF 264

Db 394 -----SVERAANYSPVLSTGGVLFN--GGTDYFRALSOETGELTMOQLATVAVS 443

Db 265 IGFDTGEVTVKAVADFDGLTGTNNLVFENLGGPTSEITQPTSMKLEVYTKRGCT--A 322

Db 444 GGAISTEVDGQVVAIAGG 462

Db 323 GDPISWTVSGTLAVTVGG 341

RESULT 6

US-09-031-655-10

Sequence 10, Application US/09031655

Patent No. 6017759

GENERAL INFORMATION:

APPLICANT: VAKHARIA, VIKRAM

APPLICANT: SNYDER, DAVID B

TITLE OF INVENTION: CHIMERIC INFECTIOUS BURSAL DISEASE VIRUS

TITLE OF INVENTION: CDNA CLONES, EXPRESSION PRODUCTS AND VACCINES BASED

NUMBER OF SEQUENCES: 15

CORRESPONDENCE ADDRESSES:

ADDRESS: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT

STREET: 1755 S. JEFFERSON DAVID HIGHWAY, FOURTH FLOOR

CITY: ARLINGTON

STATE: VIRGINIA

COUNTRY: USA

ZIP: 22202

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/031,655

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/219,262

FILING DATE: 29-MAR-1994

ATTORNEY/AGENT INFORMATION:

NAME: OBLON, NORMAN F

REGISTRATION NUMBER: 24,618

REFERENCE/DOCKET NUMBER: 2747-047-27

TELEPHONE: (703) 413-3000

TELEFAX: (703) 413-2220

TELEX: 248855 OPAT UR

INFORMATION FOR SEQ ID NO: 10:

SEQUENCE CHARACTERISTICS:

LENGTH: 1012 amino acids

TYPE: amino acid

STRANDEDNESS:

MOLECULE TYPE: protein

ORIGINAL SOURCE:

ORGANISM: Infectious bursal disease virus

STRAIN: OH

US-09-031-655-10

Query Match

Best Local Similarity 4.2%; Score 105; DB 3; Length 1012;

Matches 84; Conservative 45; Mismatches 130; Indels 120; Gaps 19;

Db 166 VTNLVHGSYAV-----GPAS-----ETORGPGGTLGTWTF 199

Db 1 MTNLMHTQOIVPFRSLMPTTGPASIPDDTLKHTLRSETSTYNLTGVDGSLGVTF 60

Db 200 AVRPDGTGEIWRHQTLPD---NMDOECTFEKMTV-----NVDVOPSTE 240

Db 61 PGFP--GSVYGAHYTLQNSGSGFDQ---MLTLAONLPVSYNYCRLVRSRLTVRSSTL 113

Db 241 MEGIOSINP--NAAGGERVLTGVPCKTGTWMPDAETGEFLVARDTNQMTEST----- 294

Db 114 PGSVYALNCTINAVT-----FQGSLSL-----TDYTINGLMSATANIN 152

Db 295 DENGITVNEADILKELDEYDVOCPTFLGDMPSALNDGSIYFIPLNVC----- 347

Db 153 DKIGNLVGEGYVLSLTSTLSDYVRIGD--PIPAAGLDP-----KLMAICDSSDRP 204

Db 348 -YDMANVQD-EE-----TSMVYNTSNVTKLPKPKDMI-GRIDAIDISTGRTLM- 393

Query Match

Best Local Similarity 4.2%; Score 105; DB 3; Length 1012;

Matches 84; Conservative 45; Mismatches 130; Indels 120; Gaps 19;

Db 166 VTNLVHGSYAV-----GPAS-----ETORGPGGTLGTWTF 199

Db 1 MTNLMHTQOIVPFRSLMPTTGPASIPDDTLKHTLRSETSTYNLTGVDGSLGVTF 60

Db 200 AVRPDGTGEIWRHQTLPD---NMDOECTFEKMTV-----NVDVOPSTE 240

Fri May 24 11:27:35 2002

walick-934-95.pap.ra1

Page 5

```
Db 61 PGFP--GSVGAHHTLOSNGSYQFO-----MLTLAONTLPVSYNVCRLYSRLTVSSLT 113
OY 241 MEGLOSINP--NATGERRVLGVCKCTGTMQFDEATGEFLMARDNYONMTEST----- 294
Db 114 PGVYALNGTINAVT-----FQGSISLSEL---TDYSYNGMSAFANIN 152
OY 295 DENGIVYNEDALIKELDEYEDVCPFLGGRWPSAALNPDGSIYFPLNNVC----- 347
Db 153 DKIGVNLGEVYVLSLPTSYDLSYVRLAD-PIPAAGLDP-----KLMATCSDRPR 204
OY 348 -YDMAVADQ-EF-----TSMOYNTSNVTKLPKDKMI-GRIDALDISTGRFLM- 393
Db 205 YTVTAADERYOFSSQLIPSGVKITLFTANIDALISLVSGLIFSGVYTHSEVDVITF 264
OY 394 -----SVERAANYSPLVSTGGVLEN--GTDRIYFALASOEGETLMOTRLATVAS 443
Db 265 ICFDGEVYVKAATDFGLTGTNNLVPFNLGPTSETTOPITSMKLEVVYTKRGGT--A 322
OY 444 GOAISTEVDGMOTVVALAG 462
Db 323 GDPISWTSGLTAVTVGG 341

RESULT 7
US-08-987-519-2
; Sequence 2, Application US/08987519
; Patent No. 5952216
; GENERAL INFORMATION:
; APPLICANT: Luderer, Steven
; TITLE OF INVENTION: Synthetic HPV16 Virus-Like Particles
; FILE REFERENCE: 19853
; CURRENT APPLICATION NUMBER: US/08/987,519
; CURRENT FILING DATE: 1997-12-09
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 500
; TYPE: PRT
; ORGANISM: Human Papillomavirus
; US-08-987-519-2

Query Match 4.1%; Score 103.5; DB 2; Length 500;
Best Local Similarity 21.4%; Pred. No. 0.14; Mismatches 82; Indels 79; Gaps 11;
Matches 52; Conservative 30;

OY 36 PPAGEWISYGCNDENYRHSPLQITTEYNGOLVWARGMOPGKVOTPLIHD----- 88
Db 158 PPLGHHMKCKOCTN--TP-----VOAGDCPPLRLITSVIQQDGMVDIG 200
OY 89 -GVYLIANIVAN--GVIVASTCOY-----SPFG---CFVSGHDSATGEELMRN 131
Db 201 PGAMNFALQNKSDVPLDICTCKPYDLOMAADPYGDRLEFFL-----REOMFAR 254
OY 132 YFIPRAGEGDETMGNDYEARMTGAWGOITTYDPTNLVHGSYAVGPASEYQRTPGGT 191
Db 255 HFNRRAGEVEPV-----PDLILIKSGNRTSVASSIYVTPSGS 294
OY 192 LYGTNTRFAVRPDTGEIYWRHQTLEPRN--WDOECTFEMVTVNDVQSTENEGLOSIN 248
Db 295 LVSESAQLFNKP-----YWLQKAGCHNNGICWGN--QLFVTVVDTTSTNNITLCASTV 345
OY 249 PNA 251
Db 346 TSS 348
```

```
APPLICANT: FRAZER, Ian
APPLICANT: ZHOU, Jian
TITLE OF INVENTION: RECOMBINANT PAPILLOMA VIRUS LI
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/737,336
FILING DATE: 16-JAN-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/AU95/00292
FILING DATE: 17-MAY-1995
PRIOR APPLICATION NUMBER: AU PM 5667
ATTORNEY/AGENT INFORMATION:
NAME: Saxe, Bernard D
REGISTRATION NUMBER: 28,665
REFERENCE/DOCKET NUMBER: 65064/118
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 532 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-737-336-6

Query Match 4.0%; Score 100; DB 3; Length 532;
Best Local Similarity 21.6%; Pred. No. 0.33; Mismatches 88; Indels 122; Gaps 18;
Matches 68; Conservative 37;

OY 3 PPS--LLMASAGALALLAAPAFVOTPYTDELNPPAGEWIS-----YGCNDENYH 53
Db 122 PTTORLVMACTGL-----EYGRQPLGVGSGHPLNKYDVENSGSG 164
OY 54 SPLQITTEWV-----OLOLV-----WARGMO-----PGKVOTPLIH 87
Db 165 GNPQDNRVNVGADYKOTQCMVGCAPPLGHHMKCKOCTNTPVQAGDCPPLRLITSVIO 224
OY 88 D-----GVYLIANIVAN--GVIVASTCOY-----SPFG---CFVSGHDSA 123
Db 225 DGDMDVTFEGAMNADLOTKNSDVPDICTCKPYDLOMAADPYGDRLEFFL-----TAVP 179
OY 124 TCEELMRVYFIPRAGEGDETMGNDYEARMTGAWGOITTYDPTNLVHGSYAVGPASEYQRTPGGT 315
Db 279 RKOEFARHFNRAGEV-----EVYPTLILIKSGNRTSVG- 315
OY 180 ASETORTPGGTLTYGTNTRFAVRPDTGEIYWRHQTLEPRN--WDOECTFEMVTVNDVQSTENEGLOSIN 236
Db 316 -SSIYVTPSGSLVSESAQLFNKP-----YWLQKAGCHNNGICWGN--QLFVTVVDTT 365
OY 237 PSTENEGLOSINPNA 251
Db 366 RSTNNITLCASTVTTSS 380

RESULT 9
```

US-08-987-519-1

Sequence 1, Application US/08987519
 Patent No. 5952216
 GENERAL INFORMATION:
 APPLICANT: Ludmerer, Steven
 TITLE OF INVENTION: Synthetic HPV16 Virus-Like Particles
 FILE REFERENCE: 19853
 CURRENT APPLICATION NUMBER: US/08/987,519
 NUMBER OF SEQ ID NOS: 3
 SOFTWARE: PASCSCO for Windows Version 3.0
 SEQ ID NO 1
 LENGTH: 500
 TYPE: PRT
 ORGANISM: Human Papillomavirus
 US-08-987-519-1

Query Match
 Best Local Similarity 3.9%; Score 97.5; DB 2; Length 500;
 Matches 66; Conservative 35; Mismatches 99; Indels 109; Gaps 16;

QY 3 PTS--LWASAGALALAAAPAFQVTPVDELANPAGEMIS-----YQ--QNGENY 51
 DB 90 PTTORLWACTGL-----EVGRQPLGVGVSGHPLINKYDDVENSNGT 132
 QY 52 RHPSLDTITENYV-----QQLV-----WARGMO-----PGKVDTPL 85
 DB 133 GGN--GDNRYNVGMDYKOTOLCMVGCAPLGEHMGKTKTOSNVONGCPLLELITGV 191
 QY 86 IHD-----GYMLANIVAN-----GYVAGSTCOY-----SPFGCVFSGHDSANG 125
 DB 192 IQGDWVDTGFGAMNADLQTKSDVPLDGTGCKYDYLOMAADPYGDRLEFY--LTK 249
 QY 126 EELMRYFTIPRAGEGSDDEWGNDEANMGTGAMGQITVDYPTNLVHYGSTAVGPPASETOR 185
 DB 250 EQMFAHFENRAGTYGEVY-----PDDLIVYGKNNSSVASSITYV 289
 QY 186 GTPGGILGYNTRFAVRDTEGEIWRHQTLPDN---WQOECTFEEMATINDVOPSTEME 242
 DB 290 HTPSGLVSSBAQLFNKP---YWLQKAGHNGKICGN---HLFVTVDDTRSTMT 340
 QY 243 GLQSLNPA 251
 DB 341 LCASVSKSA 349

RESULT 10

PCT-US93-10541-2
 Sequence 2, Application PC/TUS9310541
 GENERAL INFORMATION:

APPLICANT:
 TITLE OF INVENTION: TRANSGLOTAMINASE GENE
 NUMBER OF SEQUENCES: 5
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: RICHARDS, MEDLOCK & ANDREWS
 STREET: 1201 Elm Street, Suite 4500
 CITY: Dallas
 STATE: TX
 COUNTRY: USA
 ZIP: 75270-2197
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: PCT/US93/10541
 FILING DATE:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/126119
 FILING DATE: 23-SEP-1993
 PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/969702
 FILING DATE: 03-NOV-1992
 ATTORNEY/AGENT INFORMATION:
 NAME: Eugenia S. Hansen
 REGISTRATION NUMBER: 31966
 TELEPHONE/DOCKET NUMBER: 33590CIRCPCCT, OMRF B-
 TELEPHONE: 214-939-4500
 TELEFAX: 214-939-4500
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 548 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 PCT-US93-10541-2

Query Match
 Best Local Similarity 3.8%; Score 95.5; DB 5; Length 548;
 Matches 77; Conservative 48; Mismatches 148; Indels 87; Gaps 19;

QY 118 SGHDSATGEELMRYNFIPIRAGEGDETW-----GNDYEARMTGAMGQITPDPTNLVH 171
 DB 17 NGRDHHTA-DLCKREKLIVRRQ---PFWLTLHFGNRYEA-----SVDSLTFESVY 63
 QY 172 YGSTAVGPASETORCGTILGNTNPRAVRPTGEIWRHQTLPDNNDQCTFEMMYT 231
 DB 64 -----GPAHQEA-----GKARFPLRDAVEGDMTATVVDQ---QDCTISLDT 105
 QY 232 NVDVOPSTEMEBLOSINENATGERR-----VLTVCPKKTGTM--QDQDAETGEF-L 280
 DB 106 T-----PANAPGLVRLSLERASTGYOGSSVFLGHILLFPAKCPADAVYLDSEEROEYVL 161
 QY 281 WARDTNYQ---NMIESIDENGIVTNEALIKELDEYEDVCFPL--GGRDMPSEA----- 331
 DB 162 TQGGFTYQSGAKFIKNIPW--FGQFEDGIDILCILLDVPKFLKANGRCSSRPY 219
 QY 332 -----LNPDSGITFLPLNNCYDMAVDEFTSMOV---YNTSNTKLPKGMKI 378
 DB 220 VGRVYSGMVCNDDQGVLLGRMDNNYGDVSPMSWIASVDILRRKRNKGCORVKYQCVW 279
 QY 379 GRIDAIDISTGRTLMSVERAANYSPVLSTGGVLFNGGTDRTFR---ALSOETGETLM 434
 DB 280 --PAAVACTVRLCLGIPRTVYTNVNSAHONSMLT-----EYFRNEFGEIQGDKSEMIW 332

RESULT 11

US-08-216-276A-19
 Sequence 19, Application US/08216276A
 Patent No. 5595912
 GENERAL INFORMATION:

APPLICANT: VAKHARIA, VIKRAM
 TITLE OF INVENTION: SPECIFIC DNA AND RNA SEQUENCES
 TITLE OF INVENTION: ASSOCIATED WITH US IBV VARIANTS, VECTOR CARRYING DNA
 SEQUENCES, HOST CARRYING CLONED VECTOR, DEDICED AMINO ACID
 NUMBER OF SEQUENCES: 34
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: OBION, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
 ADDRESSEE: P.C.
 STREET: 1755 S. Jefferson Davis Highway, Suite 400
 CITY: Arlington
 STATE: Virginia
 COUNTRY: U.S.A.
 ZIP: 22202
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/216,276A
 FILING DATE: 23-MAR-1994
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/083,784
 FILING DATE: 28-JUN-1993
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/519,202
 FILING DATE: 04-MAY-1990
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/227,311
 FILING DATE: 02-AUG-1988
 ATTORNEY/AGENT INFORMATION:
 NAME: Kelter, Steven B.
 REGISTRATION NUMBER: 30,073
 REFERENCE/DOCKET NUMBER: 2747-054-27 CIP
 TELEPHONE: (703) 413-3000
 TELEFAX: (703) 413-2220
 TELETYPE: 248855 OPAT UR
 INFORMATION FOR SEQ ID NO: 19:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1012 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-216-276A-19

Query Match 3.8%; Score 95; DB 1; Length 1012;
 Best Local Similarity 23.8%; Pred. No. 2.9; Indels 88; Gaps 20;
 Matches 82; Conservative 43; Mismatches 132; Indels 88; Gaps 20;

178 GPAS-----ETGRTGGLTLYGNTFRFVRPDTEIYWRHQLPRD--N 219
 Db 24 GPASIPDDLEKHLRSTSYNLTGDDTSGIYVFPDGP--GIVAHHTLGSNENYK 81
 QY 220 WDECTFEWMTNDVDPSTMEGL--QINPNAATGERRVLGVPCXGTMMQPDDET- 276
 Db 82 FDQ-----MLTFNOMLPSPSYNCRVRSRLTYSST--LPGGVYALNGT--NAVTF 129
 QY 277 -GEIWARPDYONMIESI--DENGIVYNEALIKELDYEDVYDVCPTFLGGRDWPSSA 331
 Db 130 QGISLITDYTYGMLSATANNDKIGNVLDGCVTLSPISYDGLYRLD-PIPAIG 188
 QY 332 LNP-----DSG---IYFIPLNVCYDMAVADQFTSMQVYNTSNV-----T 369
 Db 169 LDPKVVATCSDSDRPVYTI-----TAADYQFSSQ--YOTGGVTTILFSANIDAIT 238
 QY 370 KLPFGKDWIGRIDAIDISTGRTLSV--ERAAANSPLYLSTGGVLFNGGTDTRY--FRAL 425
 Db 239 SLISVGLVFTSVHSLVLTGATITLIGFDGSNAVITRAVAANG--LTTGDLNLPENLV 295
 QY 426 --SQETGETIMQTRLATYAS-----GQALSYEVDCMQYVALAGG 462
 Db 296 IPINETOPTISIKLEIYTSKSGGQEGDDQMSAGSLAVTTHGG 340

RESULT 12
 US-09-356-818A-2
 Sequence 2, Application US/09356818A
 Patent No. 6190896
 GENERAL INFORMATION:
 APPLICANT: Fraaij, Bassem M.
 TITLE OF INVENTION: Active Human Cellular Transglutaminase
 NUMBER OF SEQUENCES: 10
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Bassem M. Fraaij
 STREET: 2401 N. Stat Dr.
 CITY: Stillwater
 STATE: Oklahoma
 COUNTRY: United States of America
 ZIP: 74075

COMPUTER READABLE FORM:
 MEDIUM TYPE: 3.5, 1.44 Mb High Density diskette
 COMPUTER: IBM AT Compatible
 OPERATING SYSTEM: Windows NT
 SOFTWARE: Wordperfect 8.0 (saved in ASCII format)
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/356,818A
 FILING DATE: July 19, 1999
 CLASSIFICATION:
 ATTORNEY/AGENT INFORMATION:
 NAME:
 REGISTRATION NUMBER:
 REFERENCE/DOCKET NUMBER: 2451,002
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (405)652-7252
 TELEFAX:
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 512
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-09-356-818A-2

Query Match 3.7%; Score 94.5; DB 4; Length 512;
 Best Local Similarity 21.4%; Pred. No. 1.1; Indels 87; Gaps 19;
 Matches 77; Conservative 48; Mismatches 148; Indels 87; Gaps 19;

QY 118 SGHSATGEELRMNYPFRAGEDETW-----GNDYEAHMTGANGQITTPYTNLVH 171
 Db 17 NORDHHTA-DLCREKLVYRGO--PFWTLIFEGRNEA-----SDVSLTFSVY---- 63
 QY 172 YGSIVAGASFTGRTGGLTLYGNTFRFVRPDTEIYWRHQLPRDWDQCTFEWMT 231
 Db 64 -----GPASQEA-----GTRAFPLRDAVEEDMTATVVOQ--QDCILSLT 105
 QY 232 NVDVDPSTMEGLQINPNAATGERR-----VLTGVPCKGTGM--QDAETGEF-L 280
 Db 106 T--FANAPIGLYRLSLBASITGQSSFVLGHPILFNAMCPADAVYLDSEERQEVVL 161
 QY 281 WARDTNYQ--NMIESIDENGIVYNEALIKELDYEDVYDVCPTFL--GGRDWPSSA-- 331
 Db 162 TQGSFITYGSAKFKINIPNN--FGQFEDGILDICLLIDVNPFLKNAGRDCSSRSPY 219
 QY 332 -----LNDPSGIVFIPLNVCYDMAVADQFTSMQVYNTSNVTKLPPGKDWI 378
 Db 220 VGRVVGAVNCDQSYLLGRMNNYGDGVSQSWIGSDILRRKRNHGGQRYKYGQCVV 279
 QY 379 GRIDAIDISTGRTLSVYERAAANSPLYLSTGGVLFNGGTDTRY--ALSDQETGLM 434
 Db 280 --FAAVACTVLRCLGTPRVVYTNINSAHQNSMLI-----EYFRNERGETGDKSEMIV 332

RESULT 13
 US-08-928-361B-11
 Sequence 11, Application US/08928361B
 Patent No. 6071518
 GENERAL INFORMATION:
 APPLICANT: Petersen, Carolyn
 TITLE OF INVENTION: PEPTIDES, POLYPEPTIDES, GLYCOPROTEINS,
 TITLE OF INVENTION: THEIR FUNCTIONAL MUTANTS, VARIANTS, ANALOGS AND FRAGMENTS
 TITLE OF INVENTION: FOR TREATMENT AND DETECTION/DIAGNOSIS OF CRYPTOSPORIDIUM
 TITLE OF INVENTION: SPECIES INFECTIONS
 NUMBER OF SEQUENCES: 30
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: PETERS, VERNY, JONES & BIRSA
 STREET: 365 Sherman Avenue, Suite 6
 CITY: Palo Alto
 STATE: CA
 COUNTRY: USA
 ZIP: 94306-1840
 COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/928,361B
 FILING DATE: 12-SEP-1997
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 60/026,062
 FILING DATE: 13-SEP-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: VERNY, Hana
 REGISTRATION NUMBER: 30,518
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 650-324-1677
 TELEFAX: 650-324-1678
 INFORMATION FOR SEQ ID NO: 11:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1042 amino acids
 TYPE: amino acid
 STRANDEDNESS:
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-928-361B-11

Query Match 3.7%; Score 94; DB 3; Length 1042;
 Best Local Similarity 20.2%; Pred. No. 3.9;
 Matches 131; Conservative 66; Mismatches 212; Indels 240; Gaps 33;

25 VTPVDELNPAGWISGQNE-----NYRSLPTQITENNGOL----- 67
 Db 301 IDPTTG-LPENNPTGLHINPTNNMTDSSPAGAKYKAVNSGKTDNYGILPVDITGLPK 359
 68 -----OLV-----WARGMPGKVVYTP-----LIHGV 90
 Db 360 DVSDDIPNSTGELVDSTGKPINNYTAGIVSGKRGLEPIEDENGMLPSTKLPIDGN 419
 QY 91 MYLANIVANGVIAGSTQYSPF-GCFVSGHDSATGEELMRYFTIRAGEGDETWGNDY 149
 Db 420 NQLVNPTNSTVSGSTGSTRKPGIPVNGGVVDE-----ANDQDK--GKD- 467
 QY 150 EARMWTGAMGQITTPYTNLVHGSTA--VGPASETORGTPG----- 189
 Db 468 --GLVPPNINSINKDPVNTQYNTGNIINP--ETGKVIKPSLPGSLNYPSENPQOTD 523
 QY 190 --GLVGTNTRFAVRPTGIVRHOITLPRDN--WDOCTFEEMVNTN-----VDVQ 236
 Db 524 EITGKRPVDTVGLPVDSTGELIDPATKLPISGAVAGDEILLEVINTITTDVETGLPDLR 583
 QY 237 ---PTEMEGIQS-----INP-----NATG-----ERRVLGVPCKTGTN 269
 Db 584 TGLPRDPSVGLPOLPNCITLVDPSNKKPIPGSHSGFISGSHEDPSGKPLDPNTG 643
 QY 270 WQFD-----AETGEFLMARDINQAMT--ESIDENGIVTVEADILKEIDV-----E 314
 Db 644 LPFEDSGSLINPETGDKLQSGHSTEMPVVGKQGGGIMPEQ--ILEALAKLPISNE 702
 QY 315 YDVCP-----TFL--GGRDMPASAL----- 332
 Db 703 VNISPSSDANPDRPTMTNMKISGQTFQVDGKKTIPGSASVHTALGPTQTDPTTG 762
 QY 333 ---NPDSGYITPLNNWCYDMAVDO-----EFTSMQVNTSNWTKLPKGDMLGRIDAI 384
 Db 763 LPSDSTGELPFGFNVLPDQCEQKSGVPSVSLYKEKNIIVTEAAYGL-----PV 815
 QY 385 DISTRTLMVERAANYSPVLTG-----GGVLFNGTDRTYRRA-----LSQETGE 431
 Db 816 DKRTGFPIDP--SILPFAANGELIDPISGKYFESSINGFTSGKAGSQQSSSDSEGN 870
 QY 432 TLMQT-----RLATVASGQAISEYEVGMQVVALAG-----GGV 464

Db 871 PIDEPTNMPDPTKTKLIDPESGIALDINSVSV-FATVPGTAAPKKGCV 918

RESULT 14
 US-08-928-361B-5
 Sequence 5, Application US/08928361B
 Patent No. 6071518
 GENERAL INFORMATION:
 APPLICANT: Petersen, Carolyn
 TITLE OF INVENTION: PEPTIDES, POLYPEPTIDES, GLYCOPROTEINS,
 TITLE OF INVENTION: THEIR FUNCTIONAL MUTANTS, VARIANTS, ANALOGS AND FRAGMENTS
 TITLE OF INVENTION: FOR TREATMENT AND DETECTION/DIAGNOSIS OF CRYPTOSPORIDIUM
 NUMBER OF SEQUENCES: 30
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: PETERS, VERNY, JONES & BIRSA
 STREET: 385 Sherman Avenue, Suite 6
 CITY: Palo Alto
 STATE: CA
 COUNTRY: USA
 ZIP: 94306-1840
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/928,361B
 FILING DATE: 12-SEP-1997
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 60/026,062
 FILING DATE: 13-SEP-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: VERNY, Hana
 REGISTRATION NUMBER: 30,518
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 650-324-1677
 TELEFAX: 650-324-1678
 INFORMATION FOR SEQ ID NO: 5:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1837 amino acids
 TYPE: amino acid
 STRANDEDNESS:
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-928-361B-5

Query Match 3.7%; Score 94; DB 3; Length 1837;
 Best Local Similarity 20.2%; Pred. No. 9.7;
 Matches 131; Conservative 66; Mismatches 212; Indels 240; Gaps 33;

25 VTPVDELNPAGWISGQNE-----NYRSLPTQITENNGOL----- 67
 Db 1096 IDPTTG-LPENNPTGLHINPTNNMTDSSPAGAKYKAVNSGKTDNYGILPVDITGLPK 1154
 68 -----OLV-----WARGMPGKVVYTP-----LIHGV 90
 Db 1155 DVSDDIPNSTGELVDSTGKPINNYTAGIVSGKRGLEPIEDENGMLPSTKLPIDGN 1214
 QY 91 MYLANIVANGVIAGSTQYSPF-GCFVSGHDSATGEELMRYFTIRAGEGDETWGNDY 149
 Db 1215 NQLVNPTNSTVSGSTGSTRKPGIPVNGGVVDE-----ANDQDK--GKD- 1262
 QY 150 EARMWTGAMGQITTPYTNLVHGSTA--VGPASETORGTPG----- 189
 Db 1263 --GLVPPNINSINKDPVNTQYNTGNIINP--ETGKVIKPSLPGSLNYPSENPQOTD 1316
 QY 190 --GLVGTNTRFAVRPTGIVRHOITLPRDN--WDOCTFEEMVNTN-----VDVQ 236

Fri May 24 11:27:35 2002

walick-934-95.pcp.ra1

Db 1319 EITGKRVDTVITGLPYDPSTGELIDPAKLPISGVADELITVINTTDEVYGLPIDE 1378
OY 237 ---PSTEMGLOS-----INP-----NAATG-----ERRVLYGPCCTGM 269
Db 1379 TGLPRDPSGLPOLNGLVLPDSNKKPIPGSHSGEINGTSGESHEKDPSTGKPLDPNNG 1438
OY 270 WQPD-----AETGEFLMADNTYQKMT--ESIDENGJYTVEDAILKELDV-----E 314
Db 1439 LPDEDESGSLINPETHDKLOGSHSGTFMFPVPGKPGEGEGLMTTEQ--ILEAKNLPTSNE 1497
OY 315 YDVC-----TFL--GGRDMPESAL-----332
Db 1498 VNISPRESSDAVDRPTNTWNNKISGCTPOVDGKKTIGSAAVHTALGTPOTDPTG 1557
OY 333 ---NPDGIFIPLNVCYDMAVADQ-----EFTSMDYNTSNVTKLPKGMGRIDAI 384
Db 1558 LPSDPSGLPFIQFNVLVDPQTEQIKGSVPVSLYKKNITVETAAVGL-----PV 1610
OY 385 DISGRILMSVERAANYSPVLTG-----GGVLFNGGTDYFRA-----ISOETGE 431
Db 1611 DPKTGFPIPL-----SYLPPAKNGELIDPISGKIFSGIAGISGKAGSOSKSDSEGN 1665
OY 432 TLMQF-----RLATVASGOAISYEVDGMQYVIAAG-----GGV 464
Db 1666 PIDPSTNPPYDKTKLIDPESGIAIDNSVGV-FATVPETAAPKKGV 1713

RESULT 15
US-09-105-537-2
Sequence 2, Application US/09105537A
Patent No. 6285202
GENERAL INFORMATION:
APPLICANT: Sherman, D. H.
APPLICANT: Liu, H.
APPLICANT: Xue, Y.
APPLICANT: Zhao, L.
TITLE OF INVENTION: DNA encoding methymycin and pikromycin
FILE REFERENCE: 600,438051
CURRENT APPLICATION NUMBER: US/09/105,537A
CURRENT FILING DATE: 1998-06-26
NUMBER OF SEQ ID NOS: 43
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 2
LENGTH: 5215
TYPE: PRT
ORGANISM: Streptomyces venezuelae
US-09-105-537-2

Query Match 3.7%; Score 94; DB 4; Length 5215;
Best Local Similarity 21.3%; Pred. No. 53;
Matches 114; Conservative 49; Mismatches 160; Indels 212; Gaps 29;
OY 67 LQLVWARGMOPGKYQVITPLIHGVMYIANIVANGVYAGSTOYSPFCFVSGHDSATGE 126
Db 1992 LELALAGDHGAVRVELELPLPERGAARIQV-----VSGDGESPAQ 2039
OY 127 ELMKRYFTPRAGEBDEET--WGNDYEARMTGAMQITVPYTNLVH-----YGSTAVG 178
Db 2040 RIFGVYSTDSGDTDDAPREM-----TRHYSVLGE--GDPATESDHPGTDGSAAMP 2092
OY 179 PASETORGTPG-----GTLVGTNTRFAVRPDTGELVWRHQLPDRNWDCECTFEMMY 230
Db 2093 PAAATATPLDGYDRLAELGYGP--AFQGLTG-LW-----RDGAD-----2131
OY 231 TANDVQ--PSTEME--GLOSINP-----NAATGERVLTGVPCKTGTMOQPD 273
Db 2132 TLAELIRLPAQHSAGLFGVHPALDALHPVLEGNAG--AC-----D 2175
OY 274 AETG-----EFLMARDNTON-----MIES-- 293
Db 2176 ADTDTATRIPLPAMAGVTLHAEGATALKVRITPTGPTVTLRLDTTGAPVATVESLTL 2235

OY 294 -----IDE-----NGIVTVNEDAILKELD--VEYDVCPT 320
Db 2236 RAVAKDRIGTTAGRVADALFTVWTFETGPBPAGRGAVEVELDLGDLVYLGADV 2295
OY 321 FLAGRDW-----PSALINPDGCIYFIPLNNVCYDMAVDOEFTSMDYNTSNVTKLP-- 372
Db 2296 VLRADRWTLDDGDSAA-----ARTAVRRLAIYQELSPRRDGSRLVCYVRGA 2344
OY 373 ---PGKDMIGRIDAIDISTGRILMSVERAANYSP-----VLSTGGVLFNGGTDYFRA 424
Db 2345 VAALPGEV-----TSLATG-PLWGLVRSAGSENGRFLFLIDIGEGERDGADELIRA 2397
OY 425 -----LSOETGETIMQRLATVA-----SGOAISEYVDGMQYVIAAGGV 464
Db 2398 ATAGDEPLAARDGRLL-APRIARTALSSSEDTAGADRFGPDGIVLYIGTGGL 2451

Search completed: May 24, 2002, 10:20:03
Job time: 316 sec

Fri May 24 11:27:35 2002

walick-934-95.pep.ra1

Fri May 24 11:27:37 2002

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2002, 10:15:22 ; Search time 54.76 Seconds
(without alignments)
828.235 Million cell updates/sec

Title: WALICK-934-95.PAP
Perfect score: 2530
Sequence: 1 MKPRLMASGALALAAP.....GMOYVAIAGGVSYSGLNS 472

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 segs, 96089334 residues

283138

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: 1: PIR.71:*
2: PIR.1:*
3: PIR.3:*
4: PIR.4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	427	16.9	601	2 E95863	alcohol dehydrogen
2	422.5	16.7	708	2 S52317	alcohol dehydrogen
3	421	16.6	742	2 A49340	alcohol dehydrogen
4	409	16.2	742	2 JS0326	alcohol dehydrogen
5	407.5	15.7	738	2 S14270	alcohol dehydrogen
6	396.5	15.7	626	2 J00706	methanol dehydrog
7	359.5	14.2	573	2 B85591	glucose dehydrogen
8	357	14.1	623	2 B83399	glucose dehydrogen
9	298.5	11.8	796	1 JVO107	glucose dehydrogen
10	299.5	11.8	796	2 H90644	glucose dehydrogen
11	299.5	11.8	796	2 H85495	glucose dehydrogen
12	295.5	11.7	796	2 AG0523	glucose dehydrogen
13	289	11.4	801	1 S00943	glucose dehydrogen
14	251.5	9.9	808	1 OPKEX	glucose dehydrogen
15	237	9.4	803	2 F83360	probable quinase d
16	223	8.8	809	2 B98314	hypothetical prote
17	223	8.8	809	2 A12968	glucose dehydrogen
18	216	8.5	778	2 G98221	glucose dehydrogen
19	216	8.5	778	2 A13064	hypothetical prote
20	209.5	8.3	809	2 A41378	quinase-shikimate
21	209.5	8.3	809	2 A55547	polyvinylalcohol d
22	187	7.4	639	2 JC4881	probably secreted
23	144.5	5.3	524	2 A82580	probable serine/th
24	133	5.0	839	2 D97013	hypothetical prote
25	127.5	4.9	688	2 C83339	probable hemolysin
26	124.5	4.7	2535	2 AC0304	serine/threonine p
27	119.5	4.7	407	2 H69064	hypothetical prote
28	119	4.7	424	2 T29127	

30	119	4.7	827	2 F64512	hypothetical prote
31	118.5	4.7	386	2 A82284	conserved hypothet
32	118.5	4.7	799	2 T48889	serine/threonine p
33	117.5	4.6	1197	2 D82696	hypothetical prote
34	115.5	4.6	525	2 A99953	hypothetical prote
35	115.5	4.6	1588	2 A86036	probable adhesin Z
36	115.5	4.6	1588	2 H91188	probable adhesin E
37	115	4.5	443	2 S23771	outer membrane por
38	115	4.5	1044	2 F95375	probable drug resi
39	114	4.5	796	2 T39962	coatomer complex b
40	114	4.5	1441	2 B86807	hypothetical prote
41	112.5	4.4	3972	2 S75251	hypothetical prote
42	112	4.4	380	2 C83171	conserved hypothet
43	112	4.4	3198	2 A43426	collagen alpha 2 I
44	111	4.4	1276	2 T18526	SREP cleavage act
45	111	4.4	2508	2 S61441	surface-associated

ALIGNMENTS

RESULT 1
E95863 alcohol dehydrogenase (acceptor) (EC 1.1.99.8) large subunit [imported] - Sinothizobi
C:Species: Sinothizobium meliloti
C:Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text-change 14-Sep-2001
C:Accession: E95863
R:Finan, T.M.; Weidner, S.; Wong, K.; Buhmester, J.; Chain, P.; Vorholter, F.J.; Her
Proc. Natl. Acad. Sci. U.S.A. 96, 9889-9894, 2001
A:Title: The complete sequence of the 1.683-kb PSYMB megaplasmid from the N2-fixing e
A:Reference number: A95842; MUID:21396508; PMID:11481431
A:Accession: E95863
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-601 <KUR>
A:Cross-references: GB:AL591985; PIRN:CAC48573.1; PID:g15140045; GSPDB:GN00167
A:Experimental source: Strain 1021, megaplasmid PSYMB
R:Finan, T.M.; Weidner, S.; Wong, K.; Buhmester, J.; Chain, P.; Vorholter, F.J.; Her
Proc. Natl. Acad. Sci. U.S.A. 96, 9889-9894, 2001
A:Title: The complete sequence of the 1.683-kb PSYMB megaplasmid from the N2-fixing e
A:Reference number: A95842; MUID:21396508; PMID:11481431
A:Accession: E95863
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-601 <KUR>
A:Cross-references: GB:AL591985; PIRN:CAC48573.1; PID:g15140045; GSPDB:GN00167
A:Experimental source: Strain 1021, megaplasmid PSYMB
R:Finan, T.M.; Weidner, S.; Wong, K.; Buhmester, J.; Chain, P.; Vorholter, F.J.; Her
Proc. Natl. Acad. Sci. U.S.A. 96, 9889-9894, 2001
A:Title: The complete sequence of the 1.683-kb PSYMB megaplasmid from the N2-fixing e
A:Reference number: A95842; MUID:21396508; PMID:11481431
A:Accession: E95863
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-601 <KUR>
A:Cross-references: GB:AL591985; PIRN:CAC48573.1; PID:g15140045; GSPDB:GN00167
A:Experimental source: Strain 1021, megaplasmid PSYMB
R:Finan, T.M.; Weidner, S.; Wong, K.; Buhmester, J.; Chain, P.; Vorholter, F.J.; Her
Proc. Natl. Acad. Sci. U.S.A. 96, 9889-9894, 2001
A:Title: The complete sequence of the 1.683-kb PSYMB megaplasmid from the N2-fixing e
A:Reference number: A95842; MUID:21396508; PMID:11481431
A:Accession: E95863
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-601 <KUR>
A:Cross-references: GB:AL591985; PIRN:CAC48573.1; PID:g15140045; GSPDB:GN00167
A:Experimental source: Strain 1021, megaplasmid PSYMB
R:Finan, T.M.; Weidner, S.; Wong, K.; Buhmester, J.; Chain, P.; Vorholter, F.J.; Her
Proc. Natl. Acad. Sci. U.S.A. 96, 9889-9894, 2001
A:Title: The complete sequence of the 1.683-kb PSYMB megaplasmid from the N2-fixing e
A:Reference number: A95842; MUID:21396508; PMID:11481431
A:Accession: E95863
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-601 <KUR>
A:Cross-references: GB:AL591985; PIRN:CAC48573.1; PID:g15140045; GSPDB:GN00167
A:Experimental source: Strain 1021, megaplasmid PSYMB
R:Finan, T.M.; Weidner, S.; Wong, K.; Buhmester, J.; Chain, P.; Vorholter, F.J.; Her
Proc. Natl. Acad. Sci. U.S.A. 96, 9889-9894, 2001
A:Title: The complete sequence of the 1.683-kb PSYMB megaplasmid from the N2-fixing e
A:Reference number: A95842; MUID:21396508; PMID:11481431
A:Accession: E95863
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-601 <KUR>
A:Cross-references: GB:AL591985; PIRN:CAC48573.1; PID:g15140045; GSPDB:GN00167
A:Experimental source: Strain 1021, megaplasmid PSYMB
R:Finan, T.M.; Weidner, S.; Wong, K.; Buhmester, J.; Chain, P.; Vorholter, F.J.; Her
Proc. Natl. Acad. Sci. U.S.A. 96, 9889-9894, 2001
A:Title: The complete sequence of the 1.683-kb PSYMB megaplasmid from the N2-fixing e
A:Reference number: A95842; MUID:21396508; PMID:11481431
A:Accession: E95863
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-601 <KUR>
A:Cross-references: GB:AL591985; PIRN:CAC48573.1; PID:g15140045; GSPDB:GN00167
A:Experimental source: Strain 1021, megaplasmid PSYMB
R:Finan, T.M.; Weidner, S.; Wong, K.; Buhmester, J.; Chain, P.; Vorholter, F.J.; Her
Proc. Natl. Acad. Sci. U.S.A. 96, 9889-9894, 2001
A:Title: The complete sequence of the 1.683-kb PSYMB megaplasmid from the N2-fixing e
A:Reference number: A95842; MUID:21396508; PMID:11481431
A:Accession: E95863
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-601 <KUR>
A:Cross-references: GB:AL591985; PIRN:CAC48573.1; PID:g15140045; GSPDB:GN00167
A:Experimental source: Strain 1021, megaplasmid PSYMB
R:Finan, T.M.; Weidner, S.; Wong, K.; Buhmester, J.; Chain, P.; Vorholter, F.J.; Her
Proc. Natl. Acad. Sci. U.S.A. 96, 9889-9894, 2001
A:Title: The complete sequence of the 1.683-kb PSYMB megaplasmid from the N2-fixing e
A:Reference number: A95842; MUID:21396508; PMID:11481431
A:Accession: E95863
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-601 <KUR>
A:Cross-references: GB:AL591985; PIRN:CAC48573.1; PID:g15140045; GSPDB:GN00167
A:Experimental source: Strain 1021, megaplasmid PSYMB
R:Finan, T.M.; Weidner, S.; Wong, K.; Buhmester, J.; Chain, P.; Vorholter, F.J.; Her
Proc. Natl. Acad. Sci. U.S.A. 96, 9889-9894, 2001
A:Title: The complete sequence of the 1.683-kb PSYMB megaplasmid from the N2-fixing e
A:Reference number: A95842; MUID:21396508; PMID:11481431
A:Accession: E95863
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-601 <KUR>
A:Cross-references: GB:AL591985; PIRN:CAC48573.1; PID:g15140045; GSPDB:GN00167
A:Experimental source: Strain 1021, megaplasmid PSYMB
R:Finan, T.M.; Weidner, S.; Wong, K.; Buhmester, J.; Chain, P.; Vorholter, F.J.; Her
Proc. Natl. Acad. Sci. U.S.A. 96, 9889-9894, 2001
A:Title: The complete sequence of the 1.683-kb PSYMB megaplasmid from the N2-fixing e
A:Reference number: A95842; MUID:21396508; PMID:11481431
A:Accession: E95863
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-601 <KUR>
A:Cross-references: GB:AL591985; PIRN:CAC48573.1; PID:g15140045; GSPDB:GN00167
A:Experimental source: Strain 1021, megaplasmid PSYMB
R:Finan, T.M.; Weidner, S.; Wong, K.; Buhmester, J.; Chain, P.; Vorholter, F.J.; Her
Proc. Natl. Acad. Sci. U.S.A. 96, 9889-9894, 2001
A:Title: The complete sequence of the 1.683-kb PSYMB megaplasmid from the N2-fixing e
A:Reference number: A95842; MUID:21396508; PMID:11481431
A:Accession: E95863
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-601 <KUR>
A:Cross-references: GB:AL591985; PIRN:CAC48573.1; PID:g15140045; GSPDB:GN00167
A:Experimental source: Strain 1021, megaplasmid PSYMB
R:Finan, T.M.; Weidner, S.; Wong, K.; Buhmester, J.; Chain, P.; Vorholter, F.J.; Her
Proc. Natl. Acad. Sci. U.S.A. 96, 9889-9894, 2001
A:Title: The complete sequence of the 1.683-kb PSYMB megaplasmid from the N2-fixing e
A:Reference number: A95842; MUID:21396508; PMID:11481431
A:Accession: E95863
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-601 <KUR>
A:Cross-references: GB:AL591985; PIRN:CAC48573.1; PID:g15140045; GSPDB:GN00167
A:Experimental source: Strain 1021, megaplasmid PSYMB
R:Finan, T.M.; Weidner, S.; Wong, K.; Buhmester, J.; Chain, P.; Vorholter, F.J.; Her
Proc. Natl. Acad. Sci. U.S.A. 96, 9889-9894, 2001
A:Title: The complete sequence of the 1.683-kb PSYMB megaplasmid from the N2-fixing e
A:Reference number: A95842; MUID:21396508; PMID:11481431
A:Accession: E95863
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-601 <KUR>
A:Cross-references: GB:AL591985; PIRN:CAC48573.1; PID:g15140045; GSPDB:GN00167
A:Experimental source: Strain 1021, megaplasmid PSYMB
R:Finan, T.M.; Weidner, S.; Wong, K.; Buhmester, J.; Chain, P.; Vorholter, F.J.; Her
Proc. Natl. Acad. Sci. U.S.A. 96, 9889-9894, 2001
A:Title: The complete sequence of the 1.683-kb PSYMB megaplasmid from the N2-fixing e
A:Reference number: A95842; MUID:21396508; PMID:11481431
A:Accession: E95863
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-601 <KUR>
A:Cross-references: GB:AL591985; PIRN:CAC48573.1; PID:g15140045; GSPDB:GN00167
A:Experimental source: Strain 1021, megaplasmid PSYMB
R:Finan, T.M.; Weidner, S.; Wong, K.; Buhmester, J.; Chain, P.; Vorholter, F.J.; Her
Proc. Natl. Acad. Sci. U.S.A. 96, 9889-9894, 2001
A:Title: The complete sequence of the 1.683-kb PSYMB megaplasmid from the N2-fixing e
A:Reference number: A95842; MUID:21396508; PMID:11481431
A:Accession: E95863
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-601 <KUR>
A:Cross-references: GB:AL591985; PIRN:CAC48573.1; PID:g15140045; GSPDB:GN00167
A:Experimental source: Strain 1021, megaplasmid PSYMB
R:Finan, T.M.; Weidner, S.; Wong, K.; Buhmester, J.; Chain, P.; Vorholter, F.J.; Her
Proc. Natl. Acad. Sci. U.S.A. 96, 9889-9894, 2001
A:Title: The complete sequence of the 1.683-kb PSYMB megaplasmid from the N2-fixing e
A:Reference number: A95842; MUID:21396508; PMID:11481431
A:Accession: E95863
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-601 <KUR>
A:Cross-references: GB:AL591985; PIRN:CAC48573.1; PID:g15140045; GSPDB:GN00167
A:Experimental source: Strain 1021, megaplasmid PSYMB
R:Finan, T.M.; Weidner, S.; Wong, K.; Buhmester, J.; Chain, P.; Vorholter, F.J.; Her
Proc. Natl. Acad. Sci. U.S.A. 96, 9889-9894, 2001
A:Title: The complete sequence of the 1.683-kb PSYMB megaplasmid from the N2-fixing e
A:Reference number: A95842; MUID:21396508; PMID:11481431
A:Accession: E95863
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-601 <KUR>
A:Cross-references: GB:AL591985; PIRN:CAC48573.1; PID:g15140045; GSPDB:GN00167
A:Experimental source: Strain 1021, megaplasmid PSYMB
R:Finan, T.M.; Weidner, S.; Wong, K.; Buhmester, J.; Chain, P.; Vorholter, F.J.; Her
Proc. Natl. Acad. Sci. U.S.A. 96, 9889-9894, 2001
A:Title: The complete sequence of the 1.683-kb PSYMB megaplasmid from the N2-fixing e
A:Reference number: A95842; MUID:21396508; PMID:11481431
A:Accession: E95863
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-601 <KUR>
A:Cross-references: GB:AL591985; PIRN:CAC48573.1; PID:g15140045; GSPDB:GN00167
A:Experimental source: Strain 1021, megaplasmid PSYMB
R:Finan, T.M.; Weidner, S.; Wong, K.; Buhmester, J.; Chain, P.; Vorholter, F.J.; Her
Proc. Natl. Acad. Sci. U.S.A. 96, 9889-9894, 2001
A:Title: The complete sequence of the 1.683-kb PSYMB megaplasmid from the N2-fixing e
A:Reference number: A95842; MUID:21396508; PMID:11481431
A:Accession: E95863
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-601 <KUR>
A:Cross-references: GB:AL591985; PIRN:CAC48573.1; PID:g15140045; GSPDB:GN00167
A:Experimental source: Strain 1021, megaplasmid PSYMB
R:Finan, T.M.; Weidner, S.; Wong, K.; Buhmester, J.; Chain, P.; Vorholter, F.J.; Her
Proc. Natl. Acad. Sci. U.S.A. 96, 9889-9894, 2001
A:Title: The complete sequence of the 1.683-kb PSYMB megaplasmid from the N2-fixing e
A:Reference number: A95842; MUID:21396508; PMID:11481431
A:Accession: E95863
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-601 <KUR>
A:Cross-references: GB:AL591985; PIRN:CAC48573.1; PID:g15140045; GSPDB:GN00167
A:Experimental source: Strain 1021, megaplasmid PSYMB
R:Finan, T.M.; Weidner, S.; Wong, K.; Buhmester, J.; Chain, P.; Vorholter, F.J.; Her
Proc. Natl. Acad. Sci. U.S.A. 96, 9889-9894, 2001
A:Title: The complete sequence of the 1.683-kb PSYMB megaplasmid from the N2-fixing e
A:Reference number: A95842; MUID:21396508; PMID:11481431
A:Accession: E95863
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-601 <KUR>
A:Cross-references: GB:AL591985; PIRN:CAC48573.1; PID:g15140045; GSPDB:GN00167
A:Experimental source: Strain 1021, megaplasmid PSYMB
R:Finan, T.M.; Weidner, S.; Wong, K.; Buhmester, J.; Chain, P.; Vorholter, F.J.; Her
Proc. Natl. Acad. Sci. U.S.A. 96, 9889-9894, 2001
A:Title: The complete sequence of the 1.683-kb PSYMB megaplasmid from the N2-fixing e
A:Reference number: A95842; MUID:21396508; PMID:11481431
A:Accession: E95863
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-601 <KUR>
A:Cross-references: GB:AL591985; PIRN:CAC48573.1; PID:g15140045; GSPDB:GN00167
A:Experimental source: Strain 1021, megaplasmid PSYMB
R:Finan, T.M.; Weidner, S.; Wong, K.; Buhmester, J.; Chain, P.; Vorholter, F.J.; Her
Proc. Natl. Acad. Sci. U.S.A. 96, 9889-9894, 2001
A:Title: The complete sequence of the 1.683-kb PSYMB megaplasmid from the N2-fixing e
A:Reference number: A95842; MUID:21396508; PMID:11481431
A:Accession: E95863
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-601 <KUR>
A:Cross-references: GB:AL591985; PIRN:CAC48573.1; PID:g15140045; GSPDB:GN00167
A:Experimental source: Strain 1021, megaplasmid PSYMB
R:Finan, T.M.; Weidner, S.; Wong, K.; Buhmester, J.; Chain, P.; Vorholter, F.J.; Her
Proc. Natl. Acad. Sci. U.S.A. 96, 9889-9894, 2001
A:Title: The complete sequence of the 1.683-kb PSYMB megaplasmid from the N2-fixing e
A:Reference number: A95842; MUID:21396508; PMID:11481431
A:Accession: E95863
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-601 <KUR>
A:Cross-references: GB:AL591985; PIRN:CAC48573.1; PID:g15140045; GSPDB:GN00167
A:Experimental source: Strain 1021, megaplasmid PSYMB
R:Finan, T.M.; Weidner, S.; Wong, K.; Buhmester, J.; Chain, P.; Vorholter, F.J.; Her
Proc. Natl. Acad. Sci. U.S.A. 96, 9889-9894, 2001
A:Title: The complete sequence of the 1.683-kb PSYMB megaplasmid from the N2-fixing e
A:Reference number: A95842; MUID:21396508; PMID:11481431
A:Accession: E95863
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-601 <KUR>
A:Cross-references: GB:AL591985; PIRN:CAC48573.1; PID:g15140045; GSPDB:GN00167
A:Experimental source: Strain 1021, megaplasmid PSYMB
R:Finan, T.M.; Weidner, S.; Wong, K.; Buhmester, J.; Chain, P.; Vorholter, F.J.; Her
Proc. Natl. Acad. Sci. U.S.A. 96, 9889-9894, 2001
A:Title: The complete sequence of the 1.683-kb PSYMB megaplasmid from the N2-fixing e
A:Reference number: A95842; MUID:21396508; PMID:11481431
A:Accession: E95863
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-601 <KUR>
A:Cross-references: GB:AL591985; PIRN:CAC48573.1; PID:g15140045; GSPDB:GN00167
A:Experimental source: Strain 1021, megaplasmid PSYMB
R:Finan, T.M.; Weidner, S.; Wong, K.; Buhmester, J.; Chain, P.; Vorholter, F.J.; Her
Proc. Natl. Acad. Sci. U.S.A. 96, 9889-9894, 2001
A:Title: The complete sequence of the 1.683-kb PSYMB megaplasmid from the N2-fixing e
A:Reference number: A95842; MUID:21396508; PMID:11481431
A:Accession: E95863
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-601 <KUR>
A:Cross-references: GB:AL591985; PIRN:CAC48573.1; PID:g15140045; GSPDB:GN00167
A:Experimental source: Strain 1021, megaplasmid PSYMB
R:Finan, T.M.; Weidner, S.; Wong, K.; Buhmester, J.; Chain, P.; Vorholter, F.J.; Her
Proc. Natl. Acad. Sci. U.S.A. 96, 9889-9894, 2001
A:Title: The complete sequence of the 1.683-kb PSYMB megaplasmid from the N2-fixing e
A:Reference number: A95842; MUID:21396508; PMID:11481431
A:Accession: E95863
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-601 <KUR>
A:Cross-references: GB:AL591985; PIRN:CAC48573.1; PID:g15140045; GSPDB:GN00167
A:Experimental source: Strain 1021, megaplasmid PSYMB
R:Finan, T.M.; Weidner, S.; Wong, K.; Buhmester, J.; Chain, P.; Vorholter, F.J.; Her
Proc. Natl. Acad. Sci. U.S.A. 96, 9889-9894, 2001
A:Title: The complete sequence of the 1.683-kb PSYMB megaplasmid from the N2-fixing e
A:Reference number: A95842; MUID:21396508; PMID:11481431
A:Accession: E95863
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-601 <KUR>
A:Cross-references: GB:AL591985; PIRN:CAC48573.1; PID:g15140045; GSPDB:GN00167
A:Experimental source: Strain 1021, megaplasmid PSYMB
R:Finan, T.M.; Weidner, S.; Wong, K.; Buhmester, J.; Chain, P.; Vorholter, F.J.; Her
Proc. Natl. Acad. Sci. U.S.A. 96, 9889-9894, 2001
A:Title: The complete sequence of the 1.683-kb PSYMB megaplasmid from the N2-fixing e
A:Reference number: A95842; MUID:21396508; PMID:11481431
A:Accession: E95863
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-601 <KUR>
A:Cross-references: GB:AL591985; PIRN:CAC48573.1; PID:g15140045; GSPDB:GN00167
A:Experimental source: Strain 1021, megaplasmid PSYMB
R:Finan, T.M.; Weidner, S.; Wong, K.; Buhmester, J.; Chain, P.; Vorholter, F.J.; Her
Proc. Natl. Acad. Sci. U.S.A. 96, 9889-9894, 2001
A:Title: The complete sequence of the 1.683-kb PSYMB megaplasmid from the N2-fixing e
A:Reference number: A95842; MUID:21396508; PMID:11481431
A:Accession: E95863
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-601 <KUR>
A:Cross-references: GB:AL591985; PIRN:CAC48573.1; PID:g15140045; GSPDB:GN00167
A:Experimental source: Strain 1


```

Db 247 EGD---QWRIKGGTTMGWYSYDEPENLVYTGPNSTNNPQR--PCDNM-SMTIFARD 300
QY 203 PDGTEIWRHQTLPDMMDOECFFEMAMTNDVOPSTMEGLQININATGRRVYAGV 262
Db 301 VDTGMAMKLYOMTPHDEMIDYGVNEMILLTEQDIDK-----PRKLTTHF 344
QY 263 PCKTGTMMQDADATGEFL-----WARDTNYQWME-----SIDENGIYV 301
Db 345 D-RNGCYITMDRTGELIVAEKYDPTVMKATEVYMDPKSKRGPRQVVAQYSTEQNG-- 400
QY 302 VNEDALIKELDEYDVCPTFLGGRDMPSSAALNPDGSGTYEPLNNVCYDMAVDEFTSMD 361
Db 401 --EDNWT-----GVCPAAGTDQOPPAYSKRTLEFVPTNHYCMDEFFRVSYTAGQ 452
QY 362 VYNTSNVTKLPKPKDM---IGRDAIDISTGRTLMSEVERAANYSPVLTSGGVLENGGT 418
Db 453 PYVGATLSMPP-KDSHGGMGNFIANDNEKIKMSLEPFSVMSGALATAGDVFEYGT 511
QY 419 DRYFRALSOETGETIMOTRLATVAASGQATSYEVDGMQYVIAAG-----GVSYGSLNS 472
Db 512 EGYLKAVDAATGKELVRFKTPSGVIGNWTITARESKQYVAVLSCVGMGIGLAAGLTN 570

```

RESULT 2

```

SS2317
quinohemoprotein ethanol dehydrogenase (EC 1.1.99.-) type 1 precursor - Comamonas testos
C:Species: Comamonas testos
C:Date: 08-May-1995 #sequence_revision 21-Jul-1995 #text_change 02-Jun-2000
R:Stoovogel, J.; Krayveld, D.E.; van Sluis, C.A.; Jongejan, J.A.; de Vries, S.; Duine, J.A.
A:Title: Characterization of the gene encoding quinohemoprotein ethanol dehydrogenase
A:Reference number: S62366; MUID:96184549
A:Accession: S62366
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-708 <STO1>
A:Cross-references: EMBL:X81880; NID:9663195; PIDN:CAA57464.1; PID:9663196
A:Accession: S62373
A:Status: Preliminary
A:Molecule type: Protein
A:Residues: 32-54 <STO2>
Eur. J. Biochem. 230, 899-905, 1995
A:Title: Quinohemoprotein ethanol dehydrogenase from Comamonas testos. Purification
A:Reference number: S65908; MUID:95324580
A:Accession: S65908
A:Molecule type: Protein
A:Residues: 32-50, 'X', 52-54, 477-483, 'X', 485-490 <DED>
A:Experimental source: ATCC 15667
C:Genetics:
C:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)
C:Keywords: oxidoreductase; quinoprotein
F:1-31/Domain: signal sequence #status predicted <SIG>
F:32-708/Product: quinohemoprotein ethanol dehydrogenase type 1 #status experimental <MZ>

```

```

Query Match 16.7%; Score 422.5; DB 2; Length 708;
Best Local Similarity 23.5%; Pred. No. 3.3e-23;
Matches 139; Conservative 79; Mismatches 222; Indels 151; Gaps 18;
QY 3 PISLMAAGALALAPAFVQVTPV-----DELLANPP-AGEWISYGOEN 50
Db 12 PGRMWLLAACIG-SAAFAQGTGPAQAQAAAVORVGDGFIANAARPPDMPTIGVDYAE 69
QY 51 YHSLPTIITENYGOLOIYVWARGOPK-YVTPPLIHGCVYLA----- 94
Db 70 TRYSHTDINAMANKULGILAMSYNESTRGIVATPVVDGIMYISASWSVYVAIDRTGN 129
QY 95 -----NIVANG----- 100

```

```

Db 130 RIWTVDPIDRSTGFGKCCVVRGVALMKGVYVYVGAANDGRLIALDAATGKEVWHONTPE 189
QY 101 -----VIVASTCQYSPFGCCFVSGHDSATGELMRNFIP----- 135
Db 190 GOKSLITITAPARVFKGVYIKRGAEEGVNG-YITVADNETGERKRMSPVAGPSKPF 248
QY 136 --RAGEGDEGTWENDYEARMGTGA-----WQIYDVTVTLVYIGSTAVGPASETQRTGP 188
Db 249 EDESKMAARTW--DPSGKMEAGGGGIMDSMTFDELMTMVGTVGNSPMSHKVRSRK 306
QY 189 GGLTYGTNTRFAVRDPTGEIWRHQTLPDMMDOECFFEMAMTNDVOPSTMEGLQIN 248
Db 307 GGDNLVIAIVLADPTDKYKMHYQEPGNDMDYSTQPMILADIKI----- 353
QY 249 PNAATGERVRLGVPCSKTGMQFDATGTFELARDTNNONNIESIDNG-IYTVEDAI 307
Db 354 ---AKRPRKVLHAP-KNGFEFVLDRTNKFTSAKNPVPVMAASGYDKHGRPIGI---AA 406
QY 308 LKELDEYDVCPTFLGGRDMPSSAALNPDGSGTYEPLNNVCYDMAVDEFT----- 358
Db 407 ARDSKRPQDAVPGPYGAHNMHMSRNPOTGLVYLPDAQVNVNLMDCKMEFNAGPKPO 466
QY 359 SMDYNTSNVTKL-PRGKMIGRDAIDISTGRTLMSEVERAANYSPVLTSGGVLENGGT 417
Db 467 SGTGMNTAKRFNAEPKSKPFGRLAMPVAKAAMSEVHSPWNGGTLTAGNVVFOGT 526
QY 418 TDRFRALSOETGETIMOTRLATVAASGQATSYEVDGMQYVIAAG-----GVSYG 467
Db 527 ADRRLVAVYAAATGKELMEAPGTGVVAAPTYVNDQRYVAVVGMGVY 577

```

RESULT 3

```

A49340
alcohol dehydrogenase (EC 1.1.-.-) precursor - Acetobacter pasteurianus (strain NC113
C:Species: Acetobacter pasteurianus
C:Date: 07-Apr-1994 #sequence_revision 07-Apr-1994 #text_change 20-Jun-2000
R:Takemura, H.; Kondo, K.; Horinouchi, S.; Beppu, T.
A:Title: Induction by ethanol of alcohol dehydrogenase activity in Acetobacter pasteu
A:Reference number: A49340; MUID:94042848
A:Accession: A49340
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-742 <TAK>
A:Cross-references: GB:D13893; NID:9517067; PIDN:BA440252.1; PID:9452586
C:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)
C:Keywords: alcohol metabolism; NAD; oxidoreductase

```

```

Query Match 16.6%; Score 421; DB 2; Length 742;
Best Local Similarity 24.7%; Pred. No. 4.6e-23;
Matches 150; Conservative 72; Mismatches 205; Indels 180; Gaps 20;
QY 11 AGALALLAPAFVQVTPVDELL--ANPPAGEWISYGOENBPHSLQITTEVGOLO 68
Db 23 AALPVAAYARADGGGNTGEATIHADHPENLWSYGRYSQGRYSLQIIRSNVGLK 82
QY 69 LVNARQMGQPKV-YVPLIHGCVYL-----ANT----- 96
Db 83 LAMVYILDTNRQGEATPLVVDGIMYATTNMSKMEALDAATGKLMLQYDPKVGNIADKGC 142
QY 97 ----- 96
Db 143 CPTVNRGAGYNGKVFNGTDPGLRYAADAATGKWEVNTIPADASIGKORSTYVQCAVR 202
QY 97 VANGYIVAGSTCOYSPFGC--FVSGHDSATGELMRNFIRAGEED----- 142
Db 203 VAKGLVILGN--GSEFGARGFVSADFAGTKLWRFYTVNKNNEPDAVADVNLMSKA 260
QY 143 -ETGNDYEARMWT-----GAMGOITYDPVYTNLVHYSSTAVGPASETQRTPGTGLY--- 193
Db 261 YKTVGP--KGAVNRQGGGGTVMDSLVYDPVSDLIY---LAVG-----NGSPMNKRYKSE 309

```

[illegible]

```

RESULT 5
S14270
alcohol dehydrogenase, membrane-bound (EC 1.1.-.-) 72k chain precursor - Acetobacter
C:Species: Acetobacter polyoxogenes
C:Variety: strain NB1028
C:Date: 21-Nov-1993 #sequence-revision 26-Jul-1996 #text-change 20-Jun-2000
C:Accession: S14270
R:Takemura, T.; Fukaya, M.; Takemura, H.; Teyama, K.; Okumura, H.; Kawamura, Y.; Nishiy
Biochem. Biophys. Acta 1088, 292-300, 1991
A:Title: Cloning and sequencing of the gene cluster encoding two subunits of membrane
A:Reference number: S14270; MUID:91159482
A:Accession: S14270
A:Molecule type: DNA
A:Residues: 1-738 <TAN>
A:Cross-references: GB:ID0635; NID:g216185; PIDN:BA00528.1; PID:g216186
A:Experimental source: strain NB1028 chains
C:Complex: heterodimer of 72k and 44k chains
C:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)
C:Keywords: alcohol metabolism; membrane protein; oxidoreductase
F:1-35/Domain: signal sequence #status predicted <Sig>
F:36-738/Product: alcohol dehydrogenase 72k chain #status predicted <Mat>

Query Match      16.1%;      Score 407.5;      DB 2;      Length 738;
Best Local Similarity 23.7%;      Pred. No. 4.5e-22;      Indexes 163;      Gaps 14;
               140;      Conservative 68;      Mismatches 219;

```

[illegible]

QY 186 GTPGGTLYGTNTREAVRPPDGEIYWRHQTLPDRNDNDOCTEEMAVTNDVQPTREBGLQ 245
 Db 308 SEGKGNLFLGSLVAKPEIGETVWHEFTPEPDQWDFSTDSQINTLDP----- 357
 QY 246 SINPMAATGERRVLTGPCKTGTMMQFDEATGEFLMADINOMIESID-ENGITVWE 304
 Db 358 -----NGETRVIVYHARKNGEFTYIDAKTGEFTISCKNIVYVMASSGLDPRTGRPTNP 410
 QY 305 DALKELDVYDCPTFLGSRDMPALNPDGTYFTPLNNVCYDMKAAVDOEFT-SMDVY 363
 Db 411 DALYTLTGKEMWGIPIGLDGHNPAAAFSPRTGLVYIPAOQVFLTYNQVGGFTPHDSW 470
 QY 364 NTS---NVTKLPPG-----KMTGRIDAIDISTGRFLMSVERAANYSPLVSTGGV 412
 Db 471 NGLDMNKVGIPIPSPEAKQAFVVDLKGWIVAMPPOKQAMRVYDHKGWNGGILATGDL 530
 QY 413 LFNGSTDRFRLSQTGETLMOETRLATVASGOAISYEVDGMQVYAIAGG 462
 Db 531 LFQGLANGEFHAYDATNGSDFEFHAFADSGIAMPVYTYLANSGYVAVEVG 580

RESULT 6

J00706
 N:Altemate names: methanol dehydrogenase (acceptor) (EC 1.1.99.8) alpha chain precursor - Methylobacterium
 C:Species: Methylobacterium extorquens
 C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 08-Oct-1999
 R:Accession: J00706; S07908
 R:Anderson, D.J.; Morris, C.J.; Nunn, D.N.; Anthony, C.; Lidstrom, M.E.
 A:Title: Nucleotide sequence of the Methylobacterium extorquens AM1 mnx and mnx genes
 A:Reference number: J00706; MIM:90337342
 A:Accession: J00706
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-626 <AND>
 A:Cross-references: GB:M31108; NID:9150017; PIDN:AAA23380.1; PID:9150018
 R:Experimental source: strain AM1
 R:Nunn, D.N.; Day, D.; Anthony, C.
 Biochem. J. 260, 857-862, 1989
 A:Title: The second subunit of methanol dehydrogenase of Methylobacterium extorquens AM1
 A:Reference number: S04644; MIM:89350892
 A:Accession: S07908
 A:Molecule type: Protein
 A:Residues: 28-50, 'xx', 53 <MUN>
 A:Note: the source is designated as Methylobacterium extorquens AM1
 C:Comment: This enzyme oxidizes methanol to formaldehyde.
 C:Keywords: alcohol metabolism; oxidoreductase
 F:1-27/Domain: signal sequence #status predicted <SIG>
 F:28-626/Product: alcohol dehydrogenase (acceptor) #status experimental <MAT>

Query Match

Best Local Similarity 15.7%; Score 396.5; DB 2; Length 626;
 Matches 141; Conservative 84; Mismatches 205; Indels 159; Gaps 22;
 QY 10 SAGALALLA-AP-AFAQVTPYTDLLANPPAGE-WISYQNOENRHSPLQITTEVNGO 66
 Db 7 SVSALAMALAPALASCAIYANDKIVELSKSDNMVMPKKNDSNNESDLQIKNGKVVQ 66
 QY 67 LQLVW--ARGQPGKVVQVTPLIHGVYVYL----- 93
 Db 67 LRPATYTFSTGLNGH-EGAPLVVDGKMYIHTSPNNRTALGLDDPGTILWODKPKNPPAA 125
 QY 94 -----ANIVA----- 108
 Db 126 RAVACCDLVNRGLAYWPGDCKTPALILKTQLDGNVAALNAETGETVWAKVNSDKVSTL 185
 QY 109 QYSPF-----GC-----FVSGHDSATGEELMRYVFI-----P 135
 Db 186 TIAPYVKKKXVITIGSSGAEIYGRKLTAYDVKTGEQVWRAIATGPKDKLLASDENIKNP 245

QY 136 RAGEE--GDETMGNDYEAARMGTA---NGQIYDPVTLVHGSTANGPASTORSTPGG 190
 Db 246 HYCGKGLCTGTWED---AMKIGGGMWGWYAYDGGTILVFTGTGNPAPWNETMR--PGD 300
 QY 191 TLVGTNTRFAVRPPDGEIYWRHQTLPDRNDNDOCTEEMAVTNDVQPTREBGLQ 245
 Db 301 NKW-TMTIFGRDADTEAFQYQKTPHDEMD-----YAGVNVMLSEKD----- 344
 QY 251 AATGERVLTGVPCKTGTMMQFDEATGEFLMAR---DTNVOAMISIDENGIYVNEAI 307
 Db 345 -KDKARKKLTLPDMNGVLYTLTDRDGLVSAKRLDT--VAVFASVLTGTGPRDPEY 401
 QY 308 LKELD-VEYDVCPTFLGSRDMPALNPDGTYFTPLNNVCYDMKAAVDOEFT-SMDVY 363
 Db 402 GTRMHLKMDICPSAMGYHNGHDSYDPKRELFTFGIHNHICMDPEPMLPYRAGOFFVGA 461
 QY 367 NVTKLPPGK-----DMGRIDAIDISTGRFLMSVERAANYSPLVSTGGVLFNGGTR 420
 Db 462 TLNMPGPKGDRQNYEGSLQIKAYNALITGDYKMEKMERFAVWGTMATAGDLVFTGLD 521
 QY 421 YFRALSGTGETLMOTRLATVASGOAISYEVDGMQVYAIAGG 462
 Db 522 YLKRSDTDLLMKFKIPSGAIGVPMYTHKTQYVAT-----YGVG 565

RESULT 7

S68591
 N:Altemate names: methanol dehydrogenase (EC 1.1.1.244) heavy chain - Methylobacterium methylobacter (str
 C:Species: Methylobacterium methylobacter
 C:Date: 06-Dec-1996 #sequence_revision 13-Mar-1997 #text_change 08-Oct-1999
 R:Accession: S68591
 R:Kia, Z.; Dai, W.; Zhang, Y.; White, S.A.; Boyd, G.D.; Mathews, F.S.
 J. Mol. Biol. 259, 480-501, 1996
 A:Title: Determination of the gene sequence and the three-dimensional structure at 2.
 A:Reference number: S68591; MIM:96256524
 A:Accession: S68591
 A:Status: preliminary; nucleic acid sequence not shown
 A:Molecule type: DNA
 A:Residues: 1-573 <XIA>
 A:Cross-references: EMBL:U41040; NID:91127819; PIDN:AAA83765.1; PID:91127820
 A:Note: The authors did not translate the codons for residues 1-2
 C:Keywords: oxidoreductase

Query Match

Best Local Similarity 14.2%; Score 359.5; DB 2; Length 573;
 Matches 127; Conservative 75; Mismatches 187; Indels 153; Gaps 20;
 QY 38 AGEWISTGQNGENYRHSPLQITTEVNGOLVW--ARGM----- 75
 Db 13 AGAMPATGGYVSGHNSPLAQINKSNVKNVKAAMSFSTGVNLNHEGAPLVDDMYVHSA 72
 QY 76 -----OPCK-----VOYPLIHGVNLAIVY--ANGVIVA- 104
 Db 73 FPNNTYALNMPGKIKWQHKPRKODASTKACVCDYDRLGALAGQVKKQANGHLAL 132
 QY 105 -----GSTCYSPF-----GC-----FVSGHDSATGEELM 129
 Db 133 DAKGTINMEVEVCDPKVGSSTLTQAPFVAKDYLMGCSGAELGVGAIVNAPDLKTGLK 192
 QY 130 R-----NYFIPRAGE--EGDETWANDYEAARMGTA---NGQIYDPVTL 168
 Db 193 RAFATGSDSVRLAKDENSANPHYGFGLGTITWGD--AMKIGGGMWGWYAYDGGTIL 249
 QY 169 LVHGSTAVGPASTGTGPIPGG-----TLVGTNTRFAVRPPDGEIYWRHQTLPDRND 223
 Db 250 LFYYGSGNRPAPWNETMR--PGDNKWTMTIMGRDL-----DTGAKMGYQKTPHDEMD 301
 QY 224 CTFEMATVNDVQPTREBGLQSIINPNAATGERVLTGVPCKTGTMMQFDEATGEFLMAR 283
 Db 302 GYVQNVWLTLD-----QPVV-----GKMTPLSHIDRNGILYTLNRENGNLIVAE 344

```

Query Match      14.1%: Score 357; DB 2; Length 623;
Best Local Similarity 21.7%: Pred. No. 1,9e-18; Indels 152; Gaps 19;
Matches 130: Conservative 87; Mismatches 231;
3 PSLTWSAGALAL--LAAPAFADVPYTDILLAN--PRAGEMISTGONQENYRHSPLT 57
Y      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
9 PAILLRPSLHCALFAVALGSGAALAKDYTWEDLINDKDTTGTVLOYQMGTHAQRWSPLK 68
Y      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
58 OITTEVNGQLOLWAMGAM--QPCRVQVYTPLIHDGVVYL----- 93
Y      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
69 QVANDNVFKLTLPAMSYSGFDEKORGO-ESQALVSDQVITYIYASYSRLFALDAKTGRKLTW 127
Y      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
94 -----ANVA-----NGVIV-----AGST 107
Y      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
128 YNHRLLPDDIRCCDVVNRGAALYGDKEVFFGLDASVAVLKNKTGVYWKKKKFAHAGYTT 187
Y      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
108 COYSP-----FGCF--VSGHDSATGEELMRNYFTR----- 136
Y      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
188 MGTAFITIKDQKTKSVLLIHSSGDEFGVYGRFLAPDDPTGEEIMRPFVSGHMGRLNGK 247
Y      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
137 -----AGEGDEDTMGNDYEA---RMKGA---KQIITYDDPTNLVHGSYFVAPASFT 183
Y      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
248 DSTVTGVDYKAPSMWPDRRNSPTGKVESWHSHGGAFWOSASDAETNTIYIYAGNPGPMWTW 307
Y      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
184 QRCSTPG-----TLXGNTFAFVRPDTGELIWRHQTLPDNNNDQECTEMMTVANDVOP 237
Y      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
308 ARTAKGAGNPDYDSLX-TSGGVGVDPSSGEVKMFYQHTPPADNADFSCNNELVLEFPKAND 366
Y      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
238 STEMGELQSIINP-----AATGERVYLTVGCYCKTGTLW--QFDAETGEELMARDINYQNA 290
Y      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
367 GKITYKATAHADNRNGFVYVDRSNGKLONAFPPVNDITWASHIIDLKTG-----RPVEREGO 421
Y      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
291 IESIDENGITYVNEDAILKELDEVYDVCPTFLGMDWPSALNPDGSGYIFPLNNVCYDM 350
Y      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

```

A:Accession: J01010
A:Molecule type: DNA
A:Residues: 1-58, 'L', 60-148, 'H', 150, 'KRRCHT', 157-192, 'K', 194-796 <CLE>
A:Cross-references: GB:X51323; NID:G41553; PIDN:CAA35706.1; PID:G41554
A:Experimental source: strain K12
R:Yamada, M.; Sumi, K.; Matsushita, K.; Adachi, O.; Yamada, Y.
J. Biol. Chem. 268, 12612-12617, 1993
A:Title: Topological analysis of quinoprotein glucose dehydrogenase in *Escherichia coli*
A:Reference number: A45997; MUID:93286127
A:Accession: A45997
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-20 <IAP>
R:Fujita, N.
submitted to the EMBL Data Library, January 1994
A:Reference number: S45181
A:Accession: S45201
A:Molecule type: DNA
A:Residues: 1-796 <FNU>
A:Cross-references: EMBL:D26562; NID:G473770; PIDN:BAA05580.1; PID:G473791
A:Experimental source: strain K-12 substrain W3110
R:Yamada, M.; Asakata, S.; Sailer, M.H.; Yamada, Y.
J. Bacteriol. 175, 566-571, 1993
A:Title: Characterization of the gcd gene from *Escherichia coli* K-12 W3110 and *regula*
A:Reference number: I41226; MUID:93123180
A:Accession: I41228
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-148, 'H', 150, 'KRRCHT', 157-192, 'K', 194-665, 'H', 667-796 <RES>
A:Cross-references: GB:D12651; NID:G216555; PIDN:BAA02174.1; PID:G216556
C:Genetics:
A:Gene: gcd
A:Map position: 3 min
C:Function:
A:Description: catalyzes the oxidation of D-glucose to D-gluconic acid by ubiquinone
A:Pathway: respiratory chain
C:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone) transmembran
C:Keywords: oxidoreductase; pyrroloquinoline-quinone; respiratory chain; transmembran
E:11-37/Domain: transmembrane #status predicted <TM2>
E:11-59/Domain: transmembrane #status predicted

F:63-81/Domain: transmembrane #status predicted <TM3>
 F:96-110/Domain: transmembrane #status predicted <TM4>
 F:120-140/Domain: transmembrane #status predicted <TM5>
 F:93,95/Binding site: ubiquinone (Acy, Asp) #status predicted
 F:466/Active site: Asp #status predicted

Query Match 11.8%; Score 299.5; DB 1; Length 796;
 Best Local Similarity 21.3%; Pred. No. 4,7e-14;
 Matches 150; Conservative 68; Mismatches 188; Indels 297; Gaps 30;

```

OY 6 LMAS-----AGALALAAFAQVTVTDELLANPAGEMISYGOENRHSPLTQ 58
DB 135 LTMAGNDPQRLNGTSLADATPAEA-ISPVADQ-----DMPAIGRNGEQRESPLKQ 185
OY 59 ITTENVGQLOLWMA-----RGMOPGV--QVPLHDSQVMTLANT-----96
DB 186 INADVNHLKEAWVFRTGDVKQPNDPGEITNEVTFIKVGDLYLCTAHORLFDALDAASGK 245
OY 97 -----VANGVYA-----104
DB 246 EKMHYDPELKTNESFOHYTCRGVSYHEAKAETASPEVMADCPRIILPVNDGRLAIINAE 305
OY 105 -GSTCO-----YSP-----FGCFV 117
DB 306 NGKLCETFPANKVNLQSNMPTKPGLYEPTSPITITDKTYMAGSVTDFSTRETSGLV 365
OY 118 SGHDSATGEELMKNYFIPIRA-----GEGDETWGNDYEARMWTGANGOLITDPTVNLV 170
DB 366 RGFVNTGELLMA--FDGAKDPNALPSDEHTFTENS-----PNSWAPAAADALDLY 416
OY 171 HGSTAVGPASETRGCTPGGLTGTNTFAVRPDGTGEIYWRHOTLPBNDQECTFEEMV 230
DB 417 YLPMGVTTPDLMGKNETPEQERYASSI-LALNATTKLMSYOTVHDLMDM-----468
OY 231 TNDVOPSTEMEGLOSINPNAATGERVLTGVPCSTGMNOFDATGEEL-----280
DB 469 --LPAOPT-----LADITVN--GQKVPYIAPAKTGNIFVLDNRNCELVPAPERKVPQ 518
OY 281 WARDNTYQNMIESIDENGI-----VTNEDALIKEL--DVEYD-----316
DB 519 GAAGKDVYTPPOFSELSFRPTKDLGADMMGATMFDQLVCRVMPHQMRTEGFTPPSEQ 578
OY 317 --VCFPLFGDWPSPALNP-----PFLS-----PFLGPCOPAMQVYSALDLKT 679
DB 579 GTLVFPGNLGMEFGISVDNPREVALIANMALPVSCLKLPKPGCNPEOPKDAKGTGE 638
OY 336 SGI-----YFIPLNNVCYDMAVDQETSMQVNTISVTKLPKGMIGRIDALDIST 388
DB 639 SGIOPOYGVPRGYTLN-----PFLS-----PFLGPCOPAMQVYSALDLKT 679
OY 389 GRTLSVERAANYS-----PV-----LSTGGVLFNGCT--DRYFRLASQE 428
DB 680 NEVYMKRIGTPODSMPFPMPVPFNGMGMPLGGPISTAGNVLFIAATADNYLRAVYNS 739
OY 429 TGETTLMOTRLATVASGA--ISYEVDGMQYVAI-AGGCVSYGS 468
DB 740 NGEKLMOGRLP--AGGQATPMTEYVNGKQYVVISAGHGSGFGT 780

```

RESULT 10
 H90644
 glucose dehydrogenase [imported] - Escherichia coli (strain O157:H7, substrain RMD 0505)
 C:Species: Escherichia coli
 C>Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Aug-2001
 C:Accession: H90644
 R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.;
 gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shingawa, H.
 A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gen
 A:Reference number: A99629; MUID:21156231; PMID:11258796
 A:Accession: H90644
 A:Status: preliminary

A:Molecule type: DNA
 A:Residues: 1-796 <HAY>
 A:Cross-references: GB:BA000007; PIDN:BA833551.1; PID:q13359584; GSPDB:GN00154
 A:Experimental source: strain O157:H7, substrain RMD 0509952
 C:Genetics:
 A:Gene: ECS0128
 C:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)

Query Match 11.8%; Score 299.5; DB 2; Length 796;
 Best Local Similarity 21.3%; Pred. No. 4,7e-14;
 Matches 150; Conservative 68; Mismatches 188; Indels 297; Gaps 30;

```

OY 6 LMAS-----AGALALAAFAQVTVTDELLANPAGEMISYGOENRHSPLTQ 58
DB 135 LTMAGNDPQRLNGTSLADATPAEA-ISPVADQ-----DMPAIGRNGEQRESPLKQ 185
OY 59 ITTENVGQLOLWMA-----RGMOPGV--QVPLHDSQVMTLANT-----96
DB 186 INADVNHLKEAWVFRTGDVKQPNDPGEITNEVTFIKVGDLYLCTAHORLFDALDAASGK 245
OY 97 -----VANGVYA-----104
DB 246 EKMHYDPELKTNESFOHYTCRGVSYHEAKAETASPEVMADCPRIILPVNDGRLAIINAE 305
OY 105 -GSTCO-----YSP-----FGCFV 117
DB 306 NGKLCETFPANKVNLQSNMPTKPGLYEPTSPITITDKTYMAGSVTDFSTRETSGLV 365
OY 118 SGHDSATGEELMKNYFIPIRA-----GEGDETWGNDYEARMWTGANGOLITDPTVNLV 170
DB 366 RGFVNTGELLMA--FDGAKDPNALPSDEHTFTENS-----PNSWAPAAADALDLY 416
OY 171 HGSTAVGPASETRGCTPGGLTGTNTFAVRPDGTGEIYWRHOTLPBNDQECTFEEMV 230
DB 417 YLPMGVTTPDLMGKNETPEQERYASSI-LALNATTKLMSYOTVHDLMDM-----468
OY 231 TNDVOPSTEMEGLOSINPNAATGERVLTGVPCSTGMNOFDATGEEL-----280
DB 469 --LPAOPT-----LADITVN--GQKVPYIAPAKTGNIFVLDNRNCELVPAPERKVPQ 518
OY 281 WARDNTYQNMIESIDENGI-----VTNEDALIKEL--DVEYD-----316
DB 519 GAAGKDVYTPPOFSELSFRPTKDLGADMMGATMFDQLVCRVMPHQMRTEGFTPPSEQ 578
OY 317 --VCFPLFGDWPSPALNP-----PFLS-----PFLGPCOPAMQVYSALDLKT 679
DB 579 GTLVFPGNLGMEFGISVDNPREVALIANMALPVSCLKLPKPGCNPEOPKDAKGTGE 638
OY 336 SGI-----YFIPLNNVCYDMAVDQETSMQVNTISVTKLPKGMIGRIDALDIST 388
DB 639 SGIOPOYGVPRGYTLN-----PFLS-----PFLGPCOPAMQVYSALDLKT 679
OY 389 GRTLSVERAANYS-----PV-----LSTGGVLFNGCT--DRYFRLASQE 428
DB 680 NEVYMKRIGTPODSMPFPMPVPFNGMGMPLGGPISTAGNVLFIAATADNYLRAVYNS 739
OY 429 TGETTLMOTRLATVASGA--ISYEVDGMQYVAI-AGGCVSYGS 468
DB 740 NGEKLMOGRLP--AGGQATPMTEYVNGKQYVVISAGHGSGFGT 780

```

RESULT 11
 H85495
 glucose dehydrogenase [imported] - Escherichia coli (strain O157:H7, substrain EDL933)
 C:Species: Escherichia coli
 C>Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
 C:Accession: H85495
 R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; May
 Natter, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamou, K.; Apoda
 A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
 A:Reference number: A85480; MUID:21074935; PMID:11206551

Page 7

S.; Moule, S.; O'Gaora, P.
 Nature 413, 848-852, 2001
 A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens,
 A:Title: Complete genome sequence of a multiple drug resistant *Salmonella enterica* se
 A:Reference number: AB0502; PMID:11677608
 A:Accession: NC0523
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1796 <PAR>
 A:Cross-references: GB:AL513382; PIDN:CAD01327.1; PID:g16501455; GSPDB:GN00176
 C:Genetics:
 A:Gene: STY0191
 C:superfamily: glucose dehydrogenase (pyrroloquinoline quinone)

A:Gene: STY0191
C:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)

Query Match 11.7%; Score 295.5; DB:2 Length 796;
Best Local Similarity 21.5%; Pred.No.:94e-14; Indels 267; Gaps 29;
Matches 148; Conservative 66; Mismatches 207;

RESULT 13
S00943
glucose dehydrogenase (pyrroloquinoline-quinone) (EC 1.1.99.17) - *Acinetobacter calco-*
C1:Species: *Acinetobacter calcoaceticus*
C1:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999

C:Accession: S00943
 R:Cleaton-Jansen, A.M.; Goosen, N.; Odle, G.; van de Putte, P.
 Nucleic Acids Res. 16, 6228, 1988
 A:Title: Nucleotide sequence of the gene coding for quonoprotein glucose dehydrogenase
 A:Reference number: S00943; MID:88289368
 A:Accession: S00943
 A:Molecule type: DNA
 A:Residues: 1-801 <CLE>
 A:Cross-references: EMBL:X07235; NID:g38711; PIDN:CAA30222.1; PID:g38712
 A:Experimental source: strain LMD 79.41
 C:Function:
 A:Description: catalyzes the oxidation of D-glucose to D-gluconic acid by ubiquinone
 C:Superfamily: respiratory chain
 C:KeyWords: glucose dehydrogenase (pyrroloquinoline-quinone)
 F:9-35/Domain: transmembrane #status predicted <TM1>
 F:33-57/Domain: transmembrane #status predicted <TM2>
 F:61-79/Domain: transmembrane #status predicted <TM3>
 F:94-108/Domain: transmembrane #status predicted <TM4>
 F:118-137/Domain: transmembrane #status predicted <TM5>
 F:91,93/Binding site: ubiquinone (Arg, Asp) #status predicted
 F:471/Active site: Asp #status predicted

Query Match 11.4%; Score 289; DB 1; Length 801;
 Best Local Similarity 20.2%; Pred. No. 2,9e-13;
 Matches 140; Conservative 76; Mismatches 174; Indels 304; Gaps 30;

20 PAFAGVTPTDELLANPAGEMISYGNQENYRSPLETTTENNVOLOLVARGMOPGK 79
 Db 152 PETAQAVGVAE-----SMPAYGRTQAGVRSPLKQINDQNVKDLKAVMT--LRIGD 202
 QY 80 V-----QYPL-----
 Db 203 LKTNDGSETTQVTPIKIGNNFICTAHQOLIAIDPARGKEKRNPKLTKSFOHLT 262
 QY 86 -----HDGVNYLAN----- 95
 Db 263 CRGVNYDANNTTFFATSILOSKSSSTQCPKRVFVAVNADPGRACITDFGNG 322
 QY 96 -----IVANGIVAGS--TCQYS--PGCVSHDSATGEEL 128
 Db 323 QVNLOEFMPAYPGVNPSPGIYIGSVTVNGSVTDNYSNKEPSG-VIRQYDVNTKLL 381
 QY 129 WRNVEIPRAGE-----EDDETGWNDYERAMWTGAMGOITDPVNLVHGSGTAVGPA 182
 Db 382 W--VFTDGAADPNAMPDEGTTFVHNS-----PNAMAPLAYDAKLDIV-IVPTGV----- 427
 QY 183 TQRTGEGTLTG-----TNRFAVRPDGTGEIYVRHOTLPDWDQECTFEMWYN 232
 Db 428 ---GTF--DIMGDRTELKERANSMLAINASTGKLVWNFQTHHDLMDVDSOPSGLAD 482
 QY 233 VDPQSTEGEGLQSTNP--NATGERRVLTGVP-----CKTG--TIMQDDET 276
 Db 483 INKKGQIVPALYVLTGKNAFVLDLR--NGQPIVPTKEPVQYAKRGFOIKGEYSKY 540
 QY 277 GEF-----LMA-----RDTNYQNTIESIDENGIVTVNEDAIL 308
 Db 541 QEFSLNLAPODKLTDKDMGATMLDQLMCRVSEKRLNMDGITYTPSENGTL----- 592
 QY 309 KELDEYDYCTELGGRDWPSSALNPS-----GIYFI----- 341
 Db 593 -----VFPGNLGVFEMGMSVNPDRQAVANPILGLFVSLLPADENRAQTAKGAGT 644
 QY 342 -----PLNNVCYDMAVDQFTSMDYNTNSVTKLPKGMIDIGRIDADISTGRTLM--- 393
 Db 645 EQGVQPMGVY-----GVEISAFSLPLGKPCPKPANGVYAGVDLKTHEVWAKR 694
 QY 394 -----SEVERAANYSPVLSTGGGVLFNGGT--DRIYRALSQSTGETLMQTR 437
 Db 695 IGTROSLPMLPOLPAVKIGVPGGLGSGISTAGNVMEVATODNRYLRANVTNGKKLMEAR 754
 QY 438 LATVYASQA--ISYEDVGMQYVAI--AGGGSVSYGS 468

Db 755 LP--AGGQATPMTYEINCKQYVYIMAGHSGSFGT 786

RESULT 14

OPKEX

C:Species: Gluconobacter oxydans
 C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 22-May-1998
 C:Accession: S17716; S19265
 R:Cleaton-Jansen, A.M.; Dekker, S.; van de Putte, P.; Goosen, N.
 MOL. Gen. Genet. 229, 206-212, 1991
 A:Title: A single amino acid substitution changes the substrate specificity of quhop
 A:Reference number: S17716; MID:92017653
 A:Accession: S17716
 A:Molecule type: DNA
 A:Residues: 1-808 <CLE>
 A:Cross-references: EMBL:X62710
 R:Goosen, N.
 submitted to the EMBL Data Library, February 1992
 A:Reference number: S19265
 A:Accession: S19265
 A:Molecule type: DNA
 A:Residues: 1-212, 'A', 214-808 <GO>
 A:Cross-references: EMBL:X62710; NID:g58416; PID:g58417
 C:Genetics:
 A:Gene: qdh
 C:Function:
 A:Description: catalyzes the oxidation of D-glucose to D-gluconic acid by ubiquinone
 C:Superfamily: respiratory chain
 C:KeyWords: glucose dehydrogenase (pyrroloquinoline-quinone)
 F:9-28/Domain: transmembrane #status predicted <TM1>
 F:35-54/Domain: transmembrane #status predicted <TM2>
 F:60-76/Domain: transmembrane #status predicted <TM3>
 F:94-110/Domain: transmembrane #status predicted <TM4>
 F:122-138/Domain: transmembrane #status predicted <TM5>
 F:91,93/Binding site: ubiquinone (Arg, Asp) #status predicted
 F:470/Active site: Asp #status predicted

Query Match 9.9%; Score 251.5; DB 1; Length 808;
 Best Local Similarity 18.8%; Pred. No. 1.7e-10;
 Matches 132; Conservative 77; Mismatches 212; Indels 281; Gaps 24;

13 ALALLAPAPQAVPTVTDLLA-----NPPAGEMISYGNQENYRSPLETTT 61
 Db 128 AVLAFASLTDPDHDISGELPTQIANASPADPDNVPASEMHAIGRTQAGRWSPLOINA 187
 QY 62 ENVGQLOLV-----ARGMOPGV--QYPLHIDGVNYLAN----- 96
 Db 188 TVNSNLKVAMHITKDMMSNDPGEOTNEAPTEFFENNTLYMCSLHOKLFAVDATGNVK 247
 QY 97 -----VANGIYV----- 103
 Db 248 YDPAKQINPFOHLTCRGVSFHEPRANAMDSGNPAFTOCARKSILPVDGRLVEVDAD 307
 QY 104 AGSTC-----OYSPRCGY----- 117
 Db 308 TGKTCGFGFNNEIDLRPNDPYTTPGQYEFTSPVITDKLIIANSATIDNGSVKQASGA 367
 QY 118 -SGHDSATGEELW-----RNYEIPRAGEGDETWGNDVEEAKWNGANGQITTPYRLVHY 172
 Db 368 TQAFVYVYTKRVMVDASNPDPNQLPDESHPVHPNSPMSIYS-----SYDANLILYI 422
 QY 173 GSTAAGPASEFQGRP--GSTLYGNTFR-----AVRPTGEIYVRHOTLPDNDQDEC 224
 Db 423 PMGV-----GTPDQMGDRTKDSEKPAFCIVALNDTKLAFQYGVYVHNDLMDL 473
 QY 225 TFEWVNTVNDQSTEGEGLQSTNPNAATGERRVLTGVCCKCTGTMQDDETG--EFLMAR 283
 Db 474 PSEPDLVDVYQKOSTLVPAIYA-----PTKTGDIFVLDLRRTGKEIYVAP 517


```

RESULT 15
F83360
glucose dehydrogenase PA2290 [imported] - Pseudomonas aeruginosa (strain PA01)
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: F83360
C:Stover, C.K.; Pham, X.Q.; Ertwin, A.L.; Mitsuuchi, S.D.; Warrenner, P.; Hickey, M.J.; Br
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folgar, K.R.; Kas, A.; Lardi, K.; Llm
: Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A:Reference number: AB2950; MUID:20437337
A:Accession: F83360
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1,803 <5'>
A:Cross-references: GB:AE004654; GB:AE004091; NID:99948311; PIDN:AA05678.1; GSPDB:GN001
A:Experimental source: strain PA01
C:Genetics:
A:Gene: gcd, PA2290
C:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)

```

Query Match	9.48;	Score 237;	DB 2;	Length 803;
Best local Similarity	19.98;	Pred. No. 2e-09;		
Matches 148;	Conservative 68;	Mismatches 192;	Indels 334;	Gaps 31

```

QY      6 LWW-----ASAGALLAAAPFAQVTVTDELLANP-----36
      |||      |||      |||      |||
Db      103 LWMFMFRRLDADGAPAPLCTALGVAVVLAAGAAVGSO-FTNPQOIVGRIDRDSGMTST 161
      |||      |||      |||      |||
QY      37 -----PAGEWISYSGONOEYRHSPLTOITTEYNGSOLYVARGMOGKY-----Q 81
      |||      |||      |||      |||      |||      |||
Db      162 APAMPDGGWQWQYGTTEGDRKSPLOKQITTPANYSQLEEAR--IRGDLPTADDPLELTNE 219
      |||      |||      |||      |||      |||      |||
QY      82 VYPLIHDCVMYLIANIVANGVYVAGSTCQYSPFCVYSGHSDATGEELW-----129
      |||      |||      |||      |||      |||      |||
Db      220 NMPFLKVNMLAAC--TAHSKYLA-----LDPTGATIMRFPDQIOSPVGF 262
      |||      |||      |||      |||
QY      130 -----RNYFIIPRAGE-----140
      |||      |||      |||      |||
Db      263 KGFAMHCTGRGVSYDEQOYARSDYGAPPAALSERAGAVASCRRLFLPTAARLAIINA 322
      |||      |||      |||      |||
QY      141 -----GDETWNQDYE-----ARMHIGAMQOITYDEVYNL----169
      |||      |||      |||      |||
Db      323 DMGKVCEDGPGKAVNDLITAGIGPFTPGGYSTSEPAVTRMLVILIGGHVJDNESINEPSGV 382
      |||      |||      |||      |||
QY      170 -----VHVGSTA-----VGPASETORGTPG-----GTLY-----193
      |||      |||      |||      |||
Db      383 IRAFVDHDKLIVWNMDSNPNDETEPLAPGKFTYENSNNMMSLASVDEKILQGVYLLPIGNOM 442
      |||      |||      |||      |||
QY      194 -----GTNI-----RF-----AVRPDTGEIVWRHOTLPIDRWMDCECFEMAVTVNDVQPS 239
      |||      |||      |||      |||

```

Db	443	PDOMGNRTPGAEEFSGAGLVALDLNGLRMNYOTHTNDLMDM-----	VGSOPTL	493
QY	240	EMEGJOSINPNAATGERRVLTVPCRTGTMMOFDETG-----		277
Db	494	-----IDLTAQGVKRALI-APTKOGSLYULDRDRGPIVPIREVPAROGAVEGDHRA	545	
QY	278	-----EFLWA-----RDTNYOMTESIDENGIYVMEDAI	307	
Db	546	PYQARSDNLRLPRLTREDMMGSSPEFDMCLRIQFSLRYEBOYTPRPSQSL-----	598	
QY	308	LKELIDVEDVCFTEFLGCDMFSAALNDSQIYVPIPLNNVCYDMAVLDEFT---SMDVY	364	
Db	599	-----IYGNGVENMGWGSVDVDRQILTFSPNYAFAVFSQWVRPDKYPSCKREGE	649	
QY	365	TSNVTK-----LPPGKMIGRIDAIDISTGRTLSVERAAA-NISPV	405	
Db	650	TSGVOPNTGARVAUIMHPMSFTGLPCQASWMDVAGIDILTTAKVVMOHKNGTSRDNTPV	709	
QY	406	-----LSTGGVYLFNGGT-DRYFRALSOETGETIMOTRLATYAASGA--IS	448	
Db	710	PGLITGVSPMSGSIITTAGVAFSLGTLQYLRAYDVKKDKGLQMARLP--AGGQAPMS	767	
QY	449	YE-VDGMOYVAI-AGGGSVYS	468	
Db	768	YTKGDKROYLYIVAGSHGSGFT	789	

Search completed: May 24, 2002, 10:21:03
Job time: 341 sec

Fri May 24 11:27:37 2002

walick-934-95.ppt.rpr

Page 10

Fri May 24 11:27:37 2002

Walick-934-95.pep.rsp

Page 1

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2002, 10:20:22 ; Search time 28.06 Seconds
(without alignments)
651.305 Million cell updates/sec

Title: WALICK-934-95.PEP
Perfect score: 2530
Sequence: 1 MKPTSLWASAGALALAAP.....GMOYVALAGGVSYSGSLNS 472

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SWISSPROT_40.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	412	16.3	739	1	DHET_ACEU
2	409	16.2	742	1	DHET_ACEU
3	407.5	15.9	757	1	DHET_ACEU
4	407.5	15.9	757	1	DHET_ACEU
5	400	15.8	631	1	DHET_ACEU
6	396.5	15.7	626	1	DHET_ACEU
7	396.5	15.6	600	1	DHET_ACEU
8	396.5	15.6	626	1	DHET_ACEU
9	357	14.1	571	1	DHET_ACEU
10	356.5	14.1	571	1	DHET_ACEU
11	299.5	11.4	801	1	DHET_ACEU
12	289	9.9	808	1	DHET_ACEU
13	251.5	8.3	809	1	DHET_ACEU
14	209.5	7.5	790	1	DHET_ACEU
15	189.5	7.5	827	1	DHET_ACEU
16	119	4.7	799	1	DHET_ACEU
17	118.5	4.7	799	1	DHET_ACEU
18	115.5	4.6	353	1	DHET_ACEU
19	115.5	4.6	353	1	DHET_ACEU
20	115	4.5	443	1	DHET_ACEU
21	114	4.5	796	1	DHET_ACEU
22	111	4.4	1276	1	DHET_ACEU
23	110.5	4.4	3317	1	DHET_ACEU
24	108.5	4.3	3354	1	DHET_ACEU
25	108	4.3	466	1	DHET_ACEU
26	105.5	4.2	593	1	DHET_ACEU
27	104.5	4.1	1279	1	DHET_ACEU
28	102	4.0	790	1	DHET_ACEU
29	102	4.0	1645	1	DHET_ACEU
30	101	4.0	1012	1	DHET_ACEU
31	101	4.0	1012	1	DHET_ACEU
32	100	4.0	500	1	DHET_ACEU
33	100	4.0	500	1	DHET_ACEU

34	99.5	3.9	807	1	AFSK_STRGR
35	99	3.9	732	1	GCP2_RAT
36	99	3.9	1365	1	GTPS_STRPO
37	98.5	3.9	1016	1	PMR_CHLTR
38	98.5	3.9	1256	1	FINC_CHICK
39	98.5	3.9	566	1	AMY_STRGR
40	97.5	3.9	501	1	VLI_HPV11
41	97.5	3.9	746	1	FEPA_ECOLI
42	97	3.8	649	1	ACES_DROME
43	97	3.8	1355	1	RPOD_ANASP
44	97	3.8	2265	1	FINC_BOVIN
45	96	3.8	566	1	AMY_STRLM

ALIGNMENTS

RESULT ID	1	STANDARD:	PRT:	739 AA.
DHET_ACEU				
AC	044002: 007952:			
DT	15-DEC-1998 (Rel. 37, Created)			
DT	15-DEC-1998 (Rel. 37, Last sequence update)			
DT	15-JUL-1999 (Rel. 38, Last annotation update)			
DE	Alcohol dehydrogenase [acceptor] precursor (EC 1.1.99.8).			
GN	ADH.			
OS	Acetobacter europaeus.			
OC	Bacteria; Proteobacteria; alpha subdivision; Acetobacteraceae;			
OC	Glucanacetobacter.			
OK	NCBI_TaxID=33995;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=DS11 / DSM 6160;			
RA	Thurner C.A.K.;			
RL	Submitted (NOV-1996) to the EMBL/Genbank/DBJ databases.			
CC	-1- CATALYTIC ACTIVITY: A primary alcohol + acceptor = an aldehyde +			
CC	reduced acceptor.			
CC	-1- COFACTOR: POO AND HEME (BY SIMILARITY).			
CC	-1- SUBUNIT: HETEROTETRAMER (DEHYDROGENASE, CYTOCHROME, AND TWO			
CC	SMALLER UNKNOWN SUBUNIT) THAT FORMS THE ALCOHOL DEHYDROGENASE			
CC	COMPLEX (BY SIMILARITY).			
CC	-1- SUBCELLULAR LOCATION: MEMBRANE-BOUND, FACING THE PERIPLASMIC			
CC	SPACE (POTENTIAL).			
CC	-1- SIMILARITY: BELONGS TO THE BACTERIAL POO DEHYDROGENASE FAMILY.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb.sib.ch/announce/			
CC	or send an email to license@sib.sib.ch).			
DR	EMBL: X82894; CAA58066.1; -			
DR	EMBL: X09480; CAA70688.1; -			
DR	HSSP: Q924J7; IFTG.			
DR	InterPro: IPR001479; Bac_POO.			
DR	InterPro: IPR002372; Bac_POO_repeat.			
DR	InterPro: IPR000345; CytC_heme_bind.			
DR	Pfam: PF01011; Bacterial_POO_6; 6.			
DR	Pfam: PS00363; BACTERIAL_POO_1; 1.			
DR	PROSITE: PS00364; BACTERIAL_POO_2; 1.			
DR	PROSITE: PS00190; CYTOCHROME_C; 1.			
KW	Oxidoreductase; POO; Heme; Hemep; Periplasmic; Membrane; Signal.			
FT	SIGNAL			
FT	CHAIN			
FT	BINDING			
FT	BINDING			
FT	BINDING			
FT	METAL			
SQ	SEQUENCE			

Query Match 16.3%; Score 412; DB 1; Length 739;

Best Local Similarity 24.0%; Pred. No. 7, 4e-24;
Matches 140; Conservative 68; Mismatches 226; Indels 150; Gaps 14;

```

OY 10 SAGALLALAAPAFQVTPYDE-----LLANPAGEMISYGOENYRHSPLTOTTT 61
DB 17 TAGICAMALISGATATMSADGCGATGEAIHADHPGMMWMTYGRYSQRRSPLDQINR 76
OY 62 ENVOLOLVARQNGQPKQV--TPLIHDCVMYL-----A 94
DB 77 SNGNKLKLVYLDLDTNRGQEGTPLYIDGMATINMSMKAVDANGKLLMSYDPRVPG 136
OY 95 NI-----
DB 137 NIADKGGCDTVNRGAAYWNGKYFGTFDGRLLTALDAKTGLMSVNTIPEALGNORSY 196
OY 97 -----VANGYVAGSTQYSPFC--FVSGHDSATGEELMRYNFIIPRAGEDEFTWGN 147
DB 197 TYDGAERIKAGVILIGN--GSEFGARGVTAFALETKVDKMRFTFAPNKNBDHTASD 254
OY 148 D-----YEARMTGA-----NGOITYDVTNLVHYGSTAVGPASETORGPGET 191
DB 255 SYLMNKAYQWSPTCAMTRKGGCGTWMDSIVYDVALVLYGVNGSPMNYKXRSCKGD 314
OY 192 LGTITRPAVRPDTEIYVHRQTLPRDNWQECTEFEMAVTVNDVOPSTEGLOSTNPA 251
DB 315 NLFGLSIALKRETEGEYVWFQETPMDQWFTSVQOIMTLDLPINET----- 362
OY 252 ATGERVLTGVPCKTGTMMQDAETGEELMARTNQNMTESID--ENGITVNEALIKE 310
DB 363 -----RHVIVAP--KNGFFYIIDAKTGEISGKNYVYVNNASGLDPKGRIPYNDALVTL 417
OY 311 LDVEVDVCPITLGRDMPALNPDGSIYFIPLANNYDMAVDEFT--SNQVYNTS--- 366
DB 418 TGKMTYIGIPDGLGHNFAAFSPKGTGLVYIPAOQVPELYTNOVGFTPHDPSNMLGDM 477
OY 367 NVTKLPFG-----KDMGRIDAIDISTGRTMSVEAANAYSPVLTSGGVLENGGT 418
DB 478 NKVGIPIPSPAKQAFKDKLGMIVAMPDQKQAEAMRVYDHKGPMNGILATGDLFGOLA 537
OY 419 DRYFRALISQETGELMOTRLAVASQALSYEDVMQYVAAAG 462
DB 538 NGEPHATDNTGSDLFHFAADSGIAPVYTLANGKQYVAVEVG 581

RESULT 2
DHET_ACEAC STANDARD: PRT: 742 AA.
ID P18278:
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Alcohol dehydrogenase [acceptor] precursor (EC 1.1.99.8).
GN ADHA OR ADH1.
OS Acetobacter aceti.
OC Bacteria; Proteobacteria; alpha subdivision; Acetobacteraceae;
OX Acetobacter
RN NCBI_Taxid=435;
RP
RX MEDLINE=9255070; PubMed=2722742;
RA Inoue T., Sunagawa M., Mori A., Imai C., Fukuda M., Takagi M.,
RT "Cloning and sequencing of the gene encoding the 72-kilodalton
RT dehydrogenase subunit of alcohol dehydrogenase from Acetobacter
RT aceti."
RL J. Bacteriol. 171:3115-3122(1989).
RN
RP
RX 3D-STRUCTURE MODELING.
RA MEDLINE=9528964; PubMed=7772016;
RA Cozier G.E., Giles T.G., Anthony C.;
RT "The structure of the quinoprotein alcohol dehydrogenase of
RT Acetobacter aceti modelled on that of methanol dehydrogenase from
RT Methylobacterium extorquens."

```

```

RL Blochem. J. 308:375-379(1995).
CC -I- CATALYTIC ACTIVITY: A primary alcohol + acceptor = an aldehyde +
CC reduced acceptor.
CC -I- COFACTOR: POO AND HEME.
CC -I- SUBUNIT: TETRAMER OF NON IDENTICAL CHAINS (DEHYDROGENASE,
CC CYTOCHROME, AND TWO SMALLER UNKNOWN SUBUNITS).
CC -I- SIMILARITY: BELONGS TO THE BACTERIAL POO DEHYDROGENASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
DR EMBL; D90004; BAA14058.1; -.
DR PIR; J50326; J50326.
DR HSSP; Q924J7; 1FLG.
DR InterPro; IPR001479; Bac_POO.
DR InterPro; IPR002372; Bac_POO_repeat.
DR InterPro; IPR000345; CytoC_heme_bind.
DR Pfam; PF01011; Bacterial_POO_6.
DR PROSITE; PS00363; BACTERIAL_POO_1; 1.
DR PROSITE; PS00364; BACTERIAL_POO_2; 1.
DR PROSITE; PS00190; CYTOCHROME_C; 1.
DR Oxioreductase; POO; Heme; Periplasmic; Membrane; Signal.
FT SIGNAL 1 35
FT CHAIN 36 742
FT ACT_SITE 343 343
FT BINDING 649 649
FT BINDING 652 652
FT METAL 653 653
SQ SEQUENCE 742 AA; 81521 MW; 9C6C9268DAB8B25A CRC64;

Query Match 16.2%; Score 409; DB 1; Length 742;
Best Local Similarity 24.2%; Pred. No. 1.3e-23;
Matches 149; Conservative 78; Mismatches 192; Indels 196; Gaps 22;

OY 11 AGALLALAAPAFQVTPYDELT--ANPAGEMISYGOENYRHSPLTOTTTNVOQLQ 68
DB 23 AALLPYAAVPAADGCGNGTEAIHADHPENWLTSGRTYSQRRSPLDQINRSWGLK 82
OY 69 LVNARQNGQPKQV--VPLIHDCVMYL-----ANT----- 96
DB 83 LIGYTLDTNRGQEGTPLYVDGIMATINMSKMLDAATGKLLMOYDPRVGNADKGC 142
OY 97 -----
DB 143 CDTVNRGAGYWGKFWGTFDGRVLAADAKTGKRVMAVNTIPADASLGKORSTYDGA VR 202
OY 97 VANGYVAGSTQYSPFC--FVSGHDSATGEELMRYNFIIPRAGEED----- 142
DB 203 VAKGLVILIGN--GGAERFARGVSAFALETKLKMRFYTPVNNKNBPDAASNTILMKA 260
OY 143 -ETWGNDEYARMT-----GAMQITVDVPTNLVHYGSTAVGPASETORGTGCTLY-- 193
DB 261 YKTWGP--KGAMVVRGGGQVWDSLVYDPSDLY--LAVG-----NGSPWNTKRYSE 309
OY 194 --GTN-----TRFAVRPDTEIYVHRQTLPRDNWQECTEFEMAVTVNDVOPSTEGLOSI 247
DB 310 GIGSLNLFGLSIALKRETEGEYVWFQETPMDQWFTSVQOIMTLDMPK----- 358
OY 248 NPNAATGE--RRVLTVGVCPTGTMMQDAETGEELMARTNQNMTESID-----EN 297
DB 359 -----SEMRHVIVAP--KNGFFYIIDAKTGEISGKNYVYVNNASGLDPKGRIPYNDALVTL 411
OY 298 GIYTVNEDALIKELDEYDVCPTFLGGRDMPALNPDGSIYFIPLANNVC----- 348
DB 412 GLVTLNG-----KFWYGI--GGPIGAHNFAMAVSPRTLHLYIIPAIQIPFGYKNOVGF 463

```

[illegible]

QY	10	SACALLLAAAPFAQVTPVTDE-----LLANPPGEMISTVGONENRHSPLQTIT	61
Dd	17	TAGTICALISGATWASADDGGATGTGAIIHADHDHGMNMTIGRTYSDOKTSPLDQINR	76
QY	62	ENWGQLOLYWARGMOPGRKYOV--TPLIHGVAVYL	94
Dd	77	SNNGNLKLAWYIDLDTNRQEGSTPLYIDGWAKATTNSMKVAVDATGKLMSYDERVHG	136
QY	95	NI-----	96
Dd	137	NIAKGCGDIVNNGAAYWNGKVYFGTFEDRLIALDAKTGLWSVNTIIPPAELGRORSY	196
QY	97	-----VANGVIYASTOCSYSPGC--FVSGSDSATEBELMRNYIFRAGEGD-----	142
Dd	197	TVDDAPRIAGRGVLTIG--GGSEFGARGVSAFDAETCGVKMDREFEYVPNPKNEPDADS	254
QY	143	-----ETWGNDEYAERMTGA-----WGQITYPVNIVHYGSTAVGPASETOR	185
Dd	255	VLMKNAYOTWS-----PIGAMTRGGCGGTWDSLVDPADVLVLYLGNGSPRNKKYR	307
QY	186	GTPPGTLXGTMFRVAVRPDGEIYWRHQULPRDNMOECTPEMATYNVDOSPSTEMEQJ	245
Dd	308	SEGKCDNLFLGSIALKPGEIYVWHQEPMPQMOPFTSDOOIMTLDLEPI-----	357
QY	246	SINPNAATGERRVLTGYPCKTGTVMQDPAETGELMARPTNVQNMTESID-ENGIVTNE	304
Dd	358	-----NGETHRVIVHARKNGFEYIIDAKTGEBISCKNIYYNNASGIDPRTGPIYNP	410
QY	305	DAIKRELDEVYVCCTFGSGDMPSAALNDPDSGIYFIPLNNVCYDMAVDOEFT-SMVY	363
Dd	411	DALTITLTKEWGICIPDGDLGHNFAMAESPRTGLVYIPAQVPEFLTNOYGSTTHPDSM	470
QY	364	MFS--NWTKLPDG-----KMIGRIDIDISTGRTLSWEERAANYSPYSTGGCV	412
Dd	471	NLDIDMNKKVGIIPSPBAKOAFYKDILKGIYAMDQOKAEARMRDHKGPMNGGITLAVGDL	530
QY	413	LFGNGTDYRFRAISOETGETLMOUHLATVASQAISTEVDGMOYAIAGS	462
Dd	531	LFGLANGEPHAIDATNGSDLHPFADSGIAPVYTIANGKOVAVEVG	580

RESULT 4

DHET_GLUOX STANDARD; PRt: 757 AA.

AC O05542:

AD 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DR 16-OCT-2001 (Rel. 40, Last annotation update)

DE Alcohoh dehydrogenase [acceptor] precursor (EC 1.1.99.8) (G3-ADH subunit 1).

DN ADHA.

GN Glucobacter oxydans (glucobacter suboxydans).

OS Bacteria: Proteobacteria: alpha subdivision: Acetobacteraceae;

OC Glucobacter.

NCBI_TaxID=442;

ON NCBI_TaxID=442;

OX (1)

RN SEQUENCE FROM N.A., AND SEQUENCE OF 36-50.

RC STRAIN-IFO 12528;

RA MEDLINE-97208225; PubMed-9055427;

RX Kondo R., Horinouchi S.;

RT "Characterization of the genes encoding the three-component membrane bound alcohol dehydrogenase from Glucobacter suboxydans and their expression in Acetobacter pasteurianus.";

RT Appl. Environ. Microbiol. 63:1131-1138(1997).

RL A-pI-FUNCTION: DEHYDRATION OF PRIMARY ALCOHOLS (EXCEPTION: METHANOL).

CC -I- FUNCTION: (BY SIMILARITY)

CC -I- CATALYTIC ACTIVITY: A primary alcohol + acceptor = an aldehyde + reduced acceptor.

CC -I- COFACTOR: POQ AND HEME.

Db

CC -1- SUBUNIT: HETEROTRIMER (DEHYDROGENASE, CYTOCHROME AND PROTEIN
CC ADHS), THAT FORMS THE ALCOHOL DEHYDROGENASE COMPLEX.
CC -1- SUBCELLULAR LOCATION: MEMBRANE-BOUND, FACING THE PERIPLASMIC
CC SPACE (POTENTIAL).
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POQ DEHYDROGENASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: D66375; BAA19753.1; -
DR HSSP: Q924J7; IFLG.
DR InterPro: IPR001479; Bac_POO.
DR InterPro: IPR002372; Bac_POO_repeat.
DR Pfam: PF01011; Bacterial_POO; 6.
DR PROSITE: PS00363; BACTERIAL_POO_1; 1.
DR PROSITE: PS00364; BACTERIAL_POO_2; 1.
DR PROSITE: PS00190; CYTOCHROME_C; 1.
DR Oxidoreductase; POQ; Heme; Periplasmic; Membrane; Signal.
KW SIGNAL 1 34
FT CHAIN 35 757
FT MOD_RES 35 35 ALCOHOL DEHYDROGENASE [ACCEPTOR].
FT ACT_SITE 342 342 PYRROLIDONE CARBOXYLIC ACID.
FT BINDING 653 653 BASE (POTENTIAL).
FT BINDING 656 656 HEME (COVALENT) (BY SIMILARITY).
FT METAL 657 657 HEME (COVALENT) (BY SIMILARITY).
FT METAL 657 657 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
SQ SEQUENCE 757 AA; 82968 MW; 39B9F90E3B947581 CRC64;

Query Match 15.9%; Score 401.5; DB 1; Length 757;
Best Local Similarity 23.0%; Pred. No. 4.9e-23;
Matches 141; Conservative 79; Mismatches 217; Indels 175; Gaps 18;

OY 7 LMASSGALALIAA--PAFAQV---TPVDELANPPAGEMISGONQENYRHSPLTQTT 61
DB 16 LLSGAAALAFSAAPVAFVAFQEDTGFATITSSDNGHP--GMLISGHSYSORISPLDQINT 74
OY 62 ENVGQQLVVARGMQPKYOV--TPLLHDGVMTL-----A 94
DB 75 ENVGKLLAHMYDLDTNRGSGTLLVNVGMVATTNWSKMLADAATGKILMSYDPKPG 134
OY 95 NIYANG----- 100
DB 135 NIADRGCCDTVSRGAAYWNGKYFFGFDSRLALDAKTGKLVMSVYTIPEAOLGHORSY 194
OY 101 -----VIVAGSTCOYSPFGCEVSGHDSATGELWRNYFIPRAGEED----- 142
DB 195 TVDGARIRAKGVLLGNGAEFGARG--FVSAFDAETSKLDRFPTVNPENKPDGAASD 253
OY 143 -----ETWGNDYEARWMTG---AMQIITPYTNLVHGSTAVGASSETGCTGGTLL 192
DB 254 ILMKRAYPTWGNKAMKQGGGCTWDSLVDPYDVLVGLGVNGSGSPWNNKFFSECKGN 313
OY 193 YGNTRFANRPDTGEIVRHOITLPRDNMDQCTEFEMAVTVNDVOPSTEMGQGISINPMA 252
DB 314 LFLGSIVAINPDTGKYVHFOETPMDEMDYTSVQIITLMPV----- 356
OY 253 TGE--RRVLGVPCKTGTWMOFADTGEFLMARDTNQNMTESIDE--NGIVYNEALIKE 310
DB 357 NEMRHVIVAP--KNGFFYIIDAFTGKFTGKPYTENMANGLDPTVGRPNVPPALMTL 415
OY 311 LDVEYDVCPTFLGGRDPSAALNDGIVYPLNNV-----CYDMAVNDDEF 357
DB 416 TGRPLUGLPGELGHNFAAAYSPKTYIVYIPAOQIPLLYDQKGFAYRHAHWLGLDM 475
OY 358 TSMYNTVNTYKLPCKDMIGRIDALIDISTGRL--WSTERAAAYS-----PVL 407
DB 476 NKIGLFDNDPBNHAAKKD-----LKVYKGTVAWDEKMAPAFTINHKGPNNGGLLA 529

OY 408 TGGVLNGSTDRYPRALSOETGETLMOTRLATASGQASIEYDGNQYAT----- 459
DB 530 TAGNVTFQGIANGFEHAYDATNGNDLVSFRAOSHAIAPPVYITANGKQYAAVEGNGGIY 589
OY 460 --AGGVSYSGSG 469
DB 590 PLYGCVARTSG 601

RESULT 5
DHML_PARDE
ID DHML_PARDE STANDARD; PRT; 631 AA.
AC P12293;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Methanol dehydrogenase subunit 1 precursor (EC 1.1.99.8) (MDH large
DE alpha subunit) (MDH).
GN MOXE.
OS Paracoccus denitrificans.
OC Bacteria; Proteobacteria; alpha subdivision; Rhodobacter group;
OC Paracoccus
OX NCBI_TaxID=266;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 33-49.
RX MEDLINE=87307969; PubMed=3114231.
RA Harms N., de Vries G.E., Maurer K., Hoogendijk J., Stoutamer A.H.;
RT "Isolation and nucleotide sequence of the methanol dehydrogenase
RT structural gene from Paracoccus denitrificans.";
RL J. Bacteriol. 169:3969-3975(1987).
CC -1- CATALYTIC ACTIVITY: A primary alcohol + acceptor = an aldehyde +
CC reduced acceptor.
CC -1- CORFACTOR: POQ.
CC -1- SUBCELLULAR LOCATION: Periplasmic.
CC -1- MISCELLANEOUS: MDH IS THE MAJOR PROTEIN IN THE CELL DURING GROWTH
CC ON METHANOL (IN P. DENITRIFICANS MDH CONSTITUTES UP TO 15% OF THE
CC TOTAL CELL PROTEIN).
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POQ DEHYDROGENASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: M17339; AAA88366.1; -
DR HSSP: P38539; AAAH.
DR InterPro: IPR001479; Bac_POO.
DR InterPro: IPR002372; Bac_POO_repeat.
DR Pfam: PF01011; Bacterial_POO; 7.
DR PROSITE: PS00363; BACTERIAL_POO_1; 1.
DR PROSITE: PS00364; BACTERIAL_POO_2; 1.
DR Oxidoreductase; POQ; Signal; Methanol utilization; Periplasmic.
KW SIGNAL 1 32
FT CHAIN 33 631
FT DISULFID 135 136 METHANOL DEHYDROGENASE SUBUNIT 1.
FT DISULFID 418 447 BY SIMILARITY.
FT ACT_SITE 335 335 BASE (POTENTIAL).
SQ SEQUENCE 631 AA; 69799 MW; 0934DC93FEC5730B CRC64;

Query Match 15.8%; Score 400; DB 1; Length 631;
Best Local Similarity 24.3%; Pred. No. 5e-23;
Matches 146; Conservative 76; Mismatches 200; Indels 176; Gaps 23;

OY 4 TSLMASSGALAL--AAPAFQVTPVTDL--LANPPAGEMISGONQENYRHSPLTQTT 60
DB 12 SSLAAVAVMGLAVLTAPATA-----NDQVLELAKDPA--NMVWTGTDYNAQVSEKTDIN 65
OY 61 TENVGQQLVVA----- 72

```

Db 66 KENNKOLRPAWSTGVILGHCESTPLVVDKMFHTPPNTTFALDINEPKILMOKPK 125
||| ||| |||
73 -----RGMO--PKRQVQPLH-----DQVMIAN-----IYANGVIV 103
||| ||| |||
Db 126 ONPRTAVTACCDVNNGLAYWPGDDQVKPLIFRTOLDPHIYAMDGEETRMINENSDIK 185
||| ||| |||
||| ||| |||
OY 104 AGSTCOYSPF-----GC-----FVSGHDSATGEELMNYFI----- 134
||| ||| |||
Db 186 VGSITLTAIPIYIKDLVYSSGAEIVGAYDVKSGEMRFAFGPDEILLADE 245
||| ||| |||
OY 135 -----PRAGEE--GDETWNDEYARMWMTGA--WGQITPVNTLVHGSYAVGPASETOR 185
||| ||| |||
Db 246 NAFNHYGOKMLGLETWEGD--AMKIGGSTMNGWAIYDPEVDLFYSSGNNAPANNETMR 302
||| ||| |||
OY 186 GTPGGTILGTMTRFAVRDPTGEIYWRHQTLPKDNMOECTFEMATNVVQSTEMEGIQ 245
||| ||| |||
Db 303 --PGDNKM--TMAINGRELTGEAKFAVQKTPHDEM-----YAGVNMMLSEODKO 351
||| ||| |||
OY 246 SINPNAATGERRVLTGPCKTGTMMQDAETGEFLMARDTNYONMIESIDE--NGITVVA 303
||| ||| |||
Db 352 -----GOMRKILTHDRNGIYVTLDRINGDLISA-----DKMDITVNVKVEYO 394
||| ||| |||
OY 304 EDAILKEIDVEY-----DVCEPFLGGRDMPSAALNDPSGIYFLELNNVCYDMAAYD 354
||| ||| |||
Db 395 LDTGLFVRDPERGTCKMDKARDICPSANGYHNOGDSYDEPRKFMGLNHCMDWEPFM 454
||| ||| |||
OY 355 QEFTSDVYNTSNVTKLPKPKDM-----IGRDAIDISTGTSLMSVRAANYSPLYSTG 409
||| ||| |||
Db 455 LPRAGGFVFGATLWYPCPKATARAGAGQIKADAIISGEMKEMKERRSWGWTWATA 514
||| ||| |||
OY 410 GGVIFNGGDRVFRALSOETGIMOTRLATVAGSALSYEVGNOYVAILAGGVSYSG 469
||| ||| |||
Db 515 GGLTFYVTLDFGFKADSDTGILMKFKLPISGVIGHPTMYKHGDRQYVAI-----MYGVG 569
||| ||| |||

```

```

RN [4]
RP X-RAY CRYSTALLOGRAPHY (1.94 ANGSTROMS).
RX MEDLINE=95253818; PubMed=7735834; Goodwin M.G., Blake C.;
RT Ghosh M., Anthony C., Harlos K.,
RL "The refined structure of the quinoprotein methanol dehydrogenase
from Methylobacterium extorquens at 1.94 A.?"
Structure 3:177-187(1995).
CC -1- CATALYTIC ACTIVITY: A primary alcohol + acceptor = an aldehyde +
reduced acceptor.
CC -1- COPFACTOR: PQQ.
CC -1- SUBUNIT: THE HOLOENZYME MDH HAS AN ALPHA-2/BETA-2 CONFIGURATION
OF A LARGE ALPHA AND A SMALL BETA SUBUNIT.
CC -1- SUBCELLULAR LOCATION: PERIPLASMIC, BUT ASSOCIATED WITH INNER
MEMBRANE.
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL PQQ DEHYDROGENASE FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
at the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@sib-sib.ch).
CC -----
CC EMBL: M31108; AAA25380.1;
CC PIR: S07908; S07908.
CC PIR: J00706; J00706.
CC HSSP: P38539; AAAH.
CC InterPro: IPR001479; Bac_PQQ_repeat.
CC InterPro: IPR002372; Bac_PQQ_repeat.
CC PROSITE: PS00363; BACTERIAL_PQQ_1; 1.
CC PROSITE: PS00364; BACTERIAL_PQQ_2; 1.
CC Oxidoreductase; PQQ; Signal; Methanol utilization; Periplasmic.
CC SIGNAL 27
CC CHAIN 28 626 METHANOL DEHYDROGENASE SUBUNIT 1.
CC FT DISULFID 130 131
CC FT DISULFID 413 442
CC FT ACT_SITE 330 330
CC FT ACT_SITE 626 626
CC FT ACT_SITE 649880AFD2AD34C CRC64;
CC SEQUENCE 626 AA; 68434 MW; 649880AFD2AD34C CRC64;

```

```

Query Match 15.7%; Score 396.5; DB 1; Length 626;
Best Local Similarity 23.9%; Pred. No. 9, 1e-23;
Matches 141; Conservative 84; Mismatches 205; Indels 159; Gaps 22;

```



```

Query Match      15.6%      Score 395.5; DB 1; Length 600;
Best Local Similarity 24.7%      Pred No. 1e-22;
Matches 148; Conservative 175; Mismatches 204; Indels 173; Gaps 24.
QY      10 SAGALIALIAPAPAOQVTPVDILNP-----PAGMISTYGENQENYRHSPLTOITTE 62
      : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| :
Db      6 NGACIALILINSGTRA-----LANEGRARRDQALQOMALQMGDIANTYRSTLDDIND 56

```

```

      1  RESULT      8
      2  DHML_METOR
      3  ID          DHML_METOR      STANDARD;      PRT;      626 AA.
      4  AC          P15279;
      5  DT      01-APR-1990 (Rel. 14, Created)
      6  DT      01-APR-1990 (Rel. 14, Last sequence update)
      7  DT      01-MAR-1992 (Rel. 41, Last annotation update)
      8  DE      Methanol dehydrogenase subunit I precursor (EC 1.1.99.8) (MDH large
      9  DE      alpha subunit) (MDH).
     10  GN      MOXF.
     11  OS      Methylobacterium organophilum XX.
     12  OC      Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
     13  OC      Methylobacterium group; Methylobacterium.
     14  OX      NCBI_Taxid=410;
     15  RN      (1)
     16  RP      SEQUENCE FROM N.A., AND SEQUENCE OF 29-43.
     17  RC      STRAIN=ATCC 27886 / DSM 760 / NCIB 11278;
     18  RX      MEDLINE=89008094, PubMed=2459109;
     19  RA      Machlin S.M., Hanson R.S.;
     20  RT      "Nucleotide sequence and transcriptional start site of the
     21  RT      methylobacterium organophilum XX methanol dehydrogenase structural
     22  RT      gene.";
     23  RL      J. Bacteriol. 170:4739-4747(1988).
     24  CC      -1- CATALYTIC ACTIVITY: A primary alcohol + acceptor = an aldehyde +
     25  CC      reduced acceptor.
     26  CC      -1- COFACTOR: PO4
     27  CC      -1- SUBUNIT: THE HOLOENZYME MDH HAS AN ALPHA-2/BETA-2 CONFIGURATION
     28  CC      OF A LARGE ALPHA AND A SMALL BETA SUBUNIT.
     29  CC      -1- SUBCELLULAR LOCATION: PERIPLASMIC, BUT ASSOCIATED WITH INNER
     30  CC      MEMBRANE.
     31  CC      -1- SIMILARITY: BELONGS TO THE BACTERIAL PO4 DEHYDROGENASE FAMILY.
     32  CC      -----
     33  CC      THIS SWISS-PROT entry is copyright. It is produced through a collaboration
     34  CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
     35  CC      the European Bioinformatics Institute. There are no restrictions on its
     36  CC      use by non-profit institutions as long as its content is in no way
  
```


modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; M2623; AAA50289.1; -
 DR HSSP; P38539; AAAH.
 DR InterPro: IPR001479; Bac_P00.
 DR InterPro: IPR002372; Bac_P00_repeat.
 DR Pfam: PF01011; Bacterial_P00_7.
 DR PROSITE; PS00363; BACTERIAL_P00_1; 1.
 DR PROSITE; PS00364; BACTERIAL_P00_2; 1.
 DR Oxidoreductase; P00; Signal; Methanol utilization; periplasmic.
 KW SIGNAL 1 28
 FT CHAIN 29 626
 FT DISULFID 130 131
 FT DISULFID 413 442
 FT ACT_SITE 330 330
 FT SEQUENCE 626 AA; 68677 MW; 8768F6B8371E5DF6 CRC64;

Query Match 15.6%; Score 394.5; DR 1; Length 626;
 Best Local Similarity 23.9%; Pred. No. 1.3e-22; Indels 159; Gaps 22;
 Matches 141; Conservative 85; Mismatches 204;

OY 10 SAGALALLA-AP-APAOYPTVDELALNPAGE-WISYGNQENRHSPLTQITTEVQ 66
 Db 7 SYVALMLALAPALSSVAYANDKVELSKDDMMVMPGKHYSDNNYSELKOVKSNVQ 66
 OY 67 LQLVW--ARGMOPGKVQVPLTHDGVTL----- 93
 Db 67 LRAVWFSGILNGH-EGAPLVVDGKMYHTSPNNTPALDLDPPGHLMQDPKONPA 125
 OY 94 -----ANIVA-----NGVVAAGSTC 108
 Db 126 RAVACDVLNRGLAYWPGDGKTPALIKTQIDRHVVALNAETGEVWKVENSIDIKVSTL 185
 OY 109 QVSPF-----GC-----FVSGHDSATGEIMRNYEI-----P 135
 Db 186 TIAYVYKDKYIIGSSGAELGVRYLTADVKGGRVAVATGDKDLLADDFNVKNA 245
 OY 136 RAGEE--GDETFWGDYFARMYGA--WQIITYDPVNTLVHYSVAVGASFGTQRTGG 190
 Db 246 HYGQGGIATWED--ANKIGGTNWGVAYDPSTNLIFGTGNPAPMNTMR--PGD 300
 OY 191 TLYGTNFFAVRPDTGETVWRHOTLPDNDQCEMTEVNVNDVQSPSTEMEGLQSIINP 250
 Db 301 NKW-TMTIFGRDADTGEAKFGYKTPHDEWDYAGVNVAM-----PSQKQ----- 344
 OY 251 AATGERVLTGVPCKTGMQFQDAETGEFLMAR--DTYNQNMIESIDENGIYVNEDAI 307
 Db 345 -KDGTRKRLTHPDNRNGIYTLDRDGLAISANKLDDT--VNVFKTVDLKTGPVADPEY 401
 OY 308 LKELD-VEYDVCPFLGGRDMPSAALNDGSIYPIPIANNVCYMAVNDQETSMDEVYNTS 366
 Db 402 GPRMDHLADVCPASMGYHNGHDSYDKRELFPWGNHNCOMDEPMLPYRAGCGFVGA 461
 OY 367 NVTKLPPGK-----DMGRDAIDISTGRILMSVERAANYSFVLSGGGVTFNGGTR 420
 Db 462 TLNMYPGPRGDRONTBGLQIKAYVALISYKWKEMERFAWGGTLATADLVFFYETLDS 521
 OY 421 YPRALSOETGETLMQTRLATVATVSGAISYVDGMQYVIAAGGVSYS 469
 Db 522 YLAKRSDGTGLMKFKIPSGAIGYPMYTHRGKQIY-----YGVG 565

RESULT 9
 EXAA_PSEAE STANDARD; PRT: 623 AA.
 AC 092417;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Quinoprotein ethanol dehydrogenase precursor (EC 1.1.99.-) (OEDH).

GN EXAA OR PA1982.
 OS Pseudomonas aeruginosa.
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 OC Pseudomonas.
 NCBI_TaxID=287;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ATCC 17933;
 RX MEDLINE-99041560; PubMed-9826187;
 RA Diehl A., Wintzingerode F., Goerisch H.;
 RT "Quinoprotein ethanol dehydrogenase of Pseudomonas aeruginosa is a
 RT homodimer: sequence of the gene and deduced structural properties of
 RT the enzyme.";
 RL Eur. J. Biochem. 257:409-419(1998).
 [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ATCC 15692 / PA01;
 RX MEDLINE-20437337; PubMed-10984043;
 RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warrenner P.,
 RA Hickey M.J., Brinkman F.S.L., Hutnagle W.O., Kowalik D.J., Lagrou M.,
 RA Garber R.L., Goltz L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
 RA Brody R.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
 RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
 RA Reizer J., Saler M.H., Hancock R.E.W., Lory S., Olson M.V.,
 RA "Complete genome sequence of Pseudomonas aeruginosa PA01, an
 RA opportunistic pathogen.";
 RL Nature 406:959-964(2000).
 [3]
 RP SEQUENCE OF 1-52 FROM N.A.
 RC STRAIN-ATCC 17933;
 RX MEDLINE-99173751; PubMed-10075429;
 RA Schobert M., Goerisch H.;
 RT "Cytochrome c550 is an essential component of the quinoprotein ethanol
 RT oxidation system in Pseudomonas aeruginosa: cloning and sequencing of
 RT the gene encoding cytochrome c550 and an adjacent acetaldehyde
 RT dehydrogenase.";
 RL Microbiology 145:471-481(1999).
 [4]
 RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF 35-616.
 RC STRAIN-ATCC 17933;
 RX MEDLINE-20202376; PubMed-10736230;
 RA Ketel T., Diehl A., Knaute T., Stezowski J.J., Hoehne W.,
 RA Goerisch H.;
 RT "X-ray structure of the quinoprotein ethanol dehydrogenase from
 RT Pseudomonas aeruginosa: basis of substrate specificity.";
 RL J. Mol. Biol. 297:961-974(2000).
 CC - FUNCTION: OXIDIZE PRIMARY ALCOHOLS AND ALSO ACTS ON SECONDARY
 CC ALCOHOL, BUT NOT HIGHLY ACTIVE ON METHANOL.
 CC - COFACTOR: PQQ AND CALCIUM.
 CC - SUBUNIT: HOMODIMER.
 CC - SUBCELLULAR LOCATION: Periplasmic.
 CC - SIMILARITY: BELONGS TO THE BACTERIAL PQQ DEHYDROGENASE FAMILY.
 CC
 CC "THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

EMBL; AJ009858; CA008896.1; -
 EMBL; AF004624; AG05370.1; -
 EMBL; AF068264; AC79657.1; -
 PDB; 1FUG; 30-AUG-00.
 DR InterPro: IPR001479; Bac_P00.
 DR InterPro: IPR002372; Bac_P00_repeat.
 DR Pfam: PF01011; Bacterial_P00_7.
 DR PROSITE; PS00363; BACTERIAL_P00_1; FALSE_NEG.
 DR PROSITE; PS00364; BACTERIAL_P00_2; FALSE_NEG.
 KW Oxidoreductase; PQQ; Periplasmic; Signal; Calcium; 3D-structure;
 KW complete proteome.
 FT SIGNAL 1 34

FT CHAIN 35 623 QUINOPROTEIN ETHANOL DEHYDROGENASE.
 FT DISULFID 139 140
 SQ SEQUENCE 623 AA; 68123 MW; 32DE5DF20B291D6 CRC64;

Query Match 14.18; Score 357; DB 1; Length 623;
 Best Local Similarity 21.7%; Pred. No. 9.5e-20;
 Matches 130; Conservative 87; Mismatches 231; Indels 152; Gaps 19;

QY 3 PTSLMASAGALV---LAAPFAQVTPYTDLLAN--PPAGEWISYQONENRHSPLT 57
 DB 9 PAGLLRSLHCLAFVALGAGALLAKDVTMEDIANDKTGDVLYQMGTHAQRWSEPLK 68
 QY 58 QITTEANGQQLYVWANG---QPGKVQVTPLIHDSVYL----- 93
 DB 69 QVNAADVFKLTPAMSYSPFDEKORQO-ESQALVSDGVITYVYASRLFALDAKTGKRLMT 127
 QY 94 -----ANIYA-----NGVIV-----AGST 107
 DB 128 YNHRLPDIRPCDDVYNNRGAALYGKVFEGTLDASVALNKNKGKVVKKRKFADHGAGYT 187
 QY 108 CQVSP-----FCGF--VSGHDSATGEEELMRNYFIR----- 136
 DB 188 MGAFLPYVDGKGVLLIHSSGDEFGVGRFLFARPDIGELIWRPFEGHMRNGK 247
 QY 137 ---AGEBDETWGNDYEA-----RMTGA---WQIYDPTVNLVHGSTAVGFASET 183
 DB 248 DSTVTGDYKASWPDNRKSPGKVESHSAGAPWOSASAPDAETNTIIVAGNPGWNTW 307
 QY 184 ORGTGPG-----TLGTNTRFAVRPDTEIWRHQTLPDMNDQCTEFEMMTVVDQV 237
 DB 308 APTAKGPNHIDYSLY-TSGGVGVDPSSGEVKKWFYQHTNDAMDPSGNNELVLFYKAKD 366
 QY 238 STEMGQLOSLNPN-----AATGERVLTGVPCKTGTW--QDAETGEFLMARNTQNM 290
 DB 367 GKIVKATAHADRNQEFVYVDRNSGKLQMAPRVNDITWASHIDLKTG-----RPERBGQ 421
 QY 291 IESIDENGITVVEDALIKELDVEYDVCPTFLGDRWPSAALNPDGTYFPIPLNNVCYDM 350
 DB 422 RPLPEPG-----OKHRAVEVSPPLGKNNMAYSDQTGLFYPAHMKEDY 471
 QY 351 MAVDQFTSMQVNTSNVTKLPPGKMIGRIDALDISGRILMSVERAANYSPVLSTGG 410
 DB 472 WTEEVSYTKSAVLGKGFIRKMYDDHVSGLKRAMPVSGKVVWHEKHEPLMAGVLTAG 531
 QY 411 GVLFGNGDTRIRALSOETGETLMQTRLATYVASGQALSYEVDMQYVAILAGGVSQSG 470
 DB 532 NLVFTGTGDTGKAFADAKSGKELMKFQTSGISVSPITWEDDGEQYLGVT---VGTGAV 588

RESULT 10

ID DHM1_METME STANDARD; PRT; 571 AA.
 AC P38539;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Methanol dehydrogenase subunit 1 (EC 1.1.99.8) (MDH large alpha subunit) (MDH).
 OS Methylophilus methylotrophus (Bacterium W3A1).
 OC Bacteria; Proteobacteria; Beta subdivision; Methylophilus group;
 OC Methylophilus;
 OX NCBI_TaxID-17;
 RN [1]
 RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS).
 RA MEDLINE-9405969; PubMed-8241148;
 RA White S., Boyd G., Mathews F.S., Xia Z.-X., Dai W.-W., Zhang Y.-F.,
 RT "The active site structure of the calcium-containing quinoprotein
 RT methanol dehydrogenase.";
 RL Biochemistry 32:12955-12958(1993).
 RN [2]
 RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS).

RX MEDLINE-93054513; PubMed-1331050;
 RA Xia Z.-X., Dai W.-W., Xiong J.-P., Hao Z.-P., Davidson V.L., White S.,
 RA Mathews F.S.;
 RT "The three-dimensional structures of methanol dehydrogenase from two
 RT methylophilic bacteria at 2.6-A resolution.";
 RL J. Biol. Chem. 267:22289-22297(1992).
 CC -1 CATALYTIC ACTIVITY: A primary alcohol + acceptor - an aldehyde +
 CC reduced acceptor.
 CC -1 COFACTOR: TWO MOLECULES OF POQ AND TWO MOLECULES OF CALCIUM
 CC PER TETRAMER.
 CC -1 SUBUNIT: THE HOLOENZYME MDH HAS AN ALPHA-2/BETA-2 CONFIGURATION
 CC OF A LARGE ALPHA AND A SMALL BETA SUBUNIT.
 CC -1 SUBCELLULAR LOCATION: PERIPLASMIC, BUT ASSOCIATED WITH INNER
 CC MEMBRANE.
 CC -1 SIMILARITY: BELONGS TO THE BACTERIAL POQ DEHYDROGENASE FAMILY.
 DR PDB; 4AAB; 08-DEC-96.
 DR InterPro; IPR001479; Bac_POQ.
 DR InterPro; IPR002372; Bac_POQ_repeat.
 DR Pfam; PF01011; Bacterial_POQ_7.
 DR PROSITE; PS00363; BACTERIAL_POQ_1, FALSE_NEG.
 DR PROSITE; PS00364; BACTERIAL_POQ_2, 1.
 KW Calcium.
 KW Oxidoreductase; POQ; Methanol utilization; Periplasmic; 3D-structure;
 FT DISULFID 103 104
 FT ACT_SITE 379 408
 FT HELIX 2 297
 FT STRAND 11 12
 FT STRAND 14 14
 FT STRAND 17 18
 FT STRAND 21 22
 FT STRAND 26 27
 FT STRAND 34 36
 FT STRAND 37 39
 FT STRAND 41 47
 FT STRAND 59 61
 FT STRAND 62 63
 FT STRAND 64 68
 FT STRAND 71 73
 FT STRAND 75 79
 FT STRAND 80 81
 FT STRAND 83 84
 FT STRAND 86 90
 FT STRAND 96 101
 FT STRAND 103 104
 FT STRAND 112 114
 FT STRAND 115 116
 FT STRAND 121 121
 FT STRAND 123 124
 FT STRAND 126 131
 FT STRAND 132 134
 FT STRAND 137 142
 FT STRAND 146 148
 FT STRAND 149 149
 FT STRAND 151 151
 FT STRAND 157 159
 FT STRAND 160 161
 FT STRAND 162 166
 FT STRAND 168 168
 FT STRAND 170 172
 FT STRAND 173 173
 FT STRAND 177 182
 FT STRAND 183 185
 FT STRAND 188 193
 FT STRAND 198 202
 FT STRAND 203 203
 FT STRAND 205 210
 FT STRAND 212 213
 FT STRAND 219 222
 FT STRAND 223 223
 FT STRAND 226 231
 FT STRAND 241 243
 FT STRAND 244 247

BASE (POTENTIAL).

RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.,
 RT "The complete genome sequence of *Escherichia coli* K-12." ;
 RL Science 277:1453-1474(1997).
 RN [5]
 RP TOPOLOGY.
 RX MEDLINE=93286127; Pubmed=8509415;
 RA Yamada M., Sumi K., Matsushita K., Adachi O., Yamada Y.,
 RT "Topological analysis of quinoprotein glucose dehydrogenase in
 RT *Escherichia coli* and its ubiquinone-binding site." ;
 RL J. Biol. Chem. 268:12812-12817(1993).
 RN [6]
 RP 3D-STRUCTURE MODELING.
 RX MEDLINE=96128046; Pubmed=8554505;
 RA Cozier G.E., Anthony C.,
 RT "Structure of the quinoprotein glucose dehydrogenase of *Escherichia*
 RT *coli* modelled on that of methanol dehydrogenase from *Methylobacterium*
 RL *Extremus*." ;
 CC Biochem. J. 312:679-685(1995).
 CC -1- FUNCTION: GDH IS PROBABLY INVOLVED IN ENERGY CONSERVATION RATHER
 CC THAN IN SUGAR METABOLISM.
 CC -1- CATALYTIC ACTIVITY: D-glucose + acceptor = D-glucono-1,5-lactone +
 CC reduced acceptor.
 CC -1- COPRATOR: POO.
 CC -1- SUBUNIT: MONOMER.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane.
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POO DEHYDROGENASE FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: X51323; CA35706.1;
 DR EMBL: D12651; BA02174.1;
 DR EMBL: D26562; CAB20298.1;
 DR EMBL: AE000122; AAC73235.1;
 DR PIR: JY0107; JY0107.
 DR HSP: P38339; 4AAM.
 DR Ecogene: EG10369; gcd.
 DR InterPro: IPR001479; Bac_POO.
 DR InterPro: IPR002372; Bac_POO_repeat.
 DR Pfam: PF01011; Bacterial_POO_7.
 DR Prosite: PS00363; BACTERIAL_POO_1; 1.
 DR Prosite: PS00364; BACTERIAL_POO_2; 1.
 KW Oxidoreductase, POO; Transmembrane; Inner membrane; Periplasmic;
 KM Complete proteome.
 FT TRANSMEM 1 10 CYTOPLASMIC (PROBABLE).
 FT TRANSMEM 11 37 PROBABLE.
 FT TRANSMEM 38 40 PERIPLASMIC (PROBABLE).
 FT TRANSMEM 41 58 PROBABLE.
 FT TRANSMEM 59 62 PROBABLE.
 FT TRANSMEM 63 81 PROBABLE.
 FT TRANSMEM 82 95 PROBABLE.
 FT TRANSMEM 96 110 PROBABLE.
 FT TRANSMEM 111 118 PROBABLE.
 FT TRANSMEM 119 141 CYTOPLASMIC (PROBABLE).
 FT TRANSMEM 142 796 PERIPLASMIC (PROBABLE).
 FT ACT_SITE 466 466 BASE (PROBABLE).
 FT ACT_SITE 59 59 R -> L (IN REF. 1).
 FT CONFLICT 149 156 TLSADAP -> HLKRRCH (IN REF. 1 AND 2).
 FT CONFLICT 193 193 N -> K (IN REF. 1 AND 2).
 FT CONFLICT 666 666 Q -> H (IN REF. 2).
 SQ SOURCE 796 AA; 86747 MW; D9EDC705A12894E9 CRC64;

Query Match 11.88; Score 299.5; DB 1; Length 796;
 Best Local Similarity 21.38; Pred. No. 3.3e-15;
 Matches 150; Conservative 68; Mismatches 188; Indels 297; Gaps 30;

OY 6 LIMS-----AGATALLAARFAQVTPYDELLANPAGEWISYQONENRHSPLQ 58
 DB 135 LTMAGFNDPEINGTLSDATPRAE-ISPVADO-----DMPAYRNOGGREFSPLQ 185
 OY 59 ITTENVCQLOLVA-----RGMPGRV--QVPLIHQVMYLAN----- 96
 DB 186 INADVHNLEKAVFRTGQVNDPGEITNEVPVIRKGVDTLYLCTAHORLALDAASGK 245
 OY 97----- 104
 DB 246 EKHVDEBELKTESFOHTKGVSTHEAKETASPEVMADCPRIILPVNDRLAINAE 305
 OY 105 -GSTCQ-----YSP-----FGCEV 117
 DB 306 NGKLETPANKGLVNLDSNMDTPKGLYEPTSPITIDTYMAGSVTDNSTRSTGVI 365
 OY 118 SGHDSATGEELMRVYFIPRA-----GERGDTGNDYEAARMGTANGQITDPVTVNLV 170
 DB 366 RGFVNTGELIMA--FDGAKDPNALPSDEHTFENS-----PSWAPAYDAKIDL 416
 OY 171 HVGSTAVPASRTGRTGGTLGTNTFRVAPRPDGEIYVRRHQTLPNDMDQECTFEAMV 230
 DB 417 YLPMGVTPPTDIMGRTPEOERYASSI-LALNATTKLANSYOTVHHDLMDMD----- 468
 OY 231 TNVDOPSTEWEGLOSINPAATGERRVLTGPKCTGTMWOFDAETGEFL----- 280
 DB 469 -LPAQPT-----LADITVA--GQKVPVYAPAPATGNIFVLDNRRELVPAPAEKVPQ 518
 OY 281 WARTQNVQNNIESIDENGI-----VIVNEDAILKEL--DVEVD----- 316
 DB 519 GAKGKDVTPPTQFSLSFRTKDLGADMGAMTFQOLVCRWVHOMRGEITPPPSQ 578
 OY 317 ---VCFPLGGRDWSAALNP-----D 335
 DB 579 GTLVFPGNLGMFEMKGISVDNREYALANPALPEVSKLIPRPGNPMQPKDANGTGE 638
 OY 336 SGT-----YIPLANNVCYDMAVDOFTSMNVYNTKLPPEKMDIGITIDIST 388
 DB 639 SGIQPOYGVPGVTLN-----PFLS-----PFLGPKOPAMGYSISLDTK 679
 OY 389 GRTLMSVERAANYS-----PY-----LSTGGVLFNGGT-DRYFRALSOE 428
 DB 680 NEVYMKRRTIGTPODSMPFPVPPVPMKMGPKMGISAGVLFLEIATADNYLRAVMS 739
 OY 429 TGETIMORLATFVASGA--ISTEVDKQYAI-AGGVSYS 468
 DB 740 NGKLMGRLE--AGGQATPMYEVNGKQYVIVSAGHSGSGT 780
 RESULT 12
 DHGA_ACICA STANDARD; PRT; 801 AA.
 ID DHGA_ACICA
 AC P05465;
 DT 01-NOV-1988 (rel. 09, Created)
 DT 01-NOV-1988 (rel. 09, Last sequence update)
 DT 16-OCT-2001 (rel. 40, Last annotation update)
 DE Glucose dehydrogenase-A [pyrroloquinoline-quinone] precursor
 DE (EC 1.1.99.17) (Quinoprotein glucose DH) (GDH-A).
 GN GDH.
 OS Acinetobacter calcoaceticus.
 OC Bacteria; Proteobacteria; gamma subdivision; Moraxellaceae;
 OC Acinetobacter.
 OX NCBI_Taxid=471;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-LMD 79.41;
 RX MEDLINE=88289368; Pubmed=3399393;
 RA Cleton-Jansen A.-M., Goosen N., Odle G., van de Putte P.;
 RT "Nucleotide sequence of the gene coding for quinoprotein glucose
 RT dehydrogenase from *Acinetobacter calcoaceticus*." ;
 RL Nucleic Acids Res. 16:6228-6228(1988).
 CC -1- CATALYTIC ACTIVITY: D-glucose + acceptor = D-glucono-1,5-lactone +
 CC reduced acceptor.

Db	593	-----	VFPGNLGVFEWGMGSVNPDRQVAVMNPILGIPVSRLLIPADPRRAQTAAGACT	644
Oy	342	----	PLNNVCDMAAVDEFTSMDEVMTSNVTKLPGKMDIGRIDAIDISTGRITLM	393
Db	645	-----	EOGQVATPVVY-----GVEISAFSLSTLGLPCQAPMAGVAVDITKHEVWKKR	694
Oy	394	-----	SEVERAANYSFVISTGGCVFNCGT-DRYFRALSQFTGETTLMOTR	437
Db	695	-----	IGTIRDSLPHLFLQLPVAKTIGVPGIGSGSISTAGNMFVATQDNLKRAFNVTNGKKLWEAR	754
Oy	438	-----	LATVASGQA--ISTEYDGMQYAI-AGGCVSGS	468
Db	755	-----	LP--AGGQATPMTETELNGKQYVIMAGHGSGCT	786
RESULT	13			
DHG_GLUOX		STANDARD;	PRT;	808 AA.
AC	P27175;			
AD	01-AUG-1992 (Rel. 23, Created)			
DT	01-AUG-1992 (Rel. 23, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Glucose dehydrogenase [pyrroloquinoline-quinone] precursor			
DE	(EC 1.1.99.17).			
GN	GDH			
OS	Glucunobacter oxydans (Glucunobacter suboxydans).			
OC	Bacteria, Proteobacteria; alpha subdivision; Acetobacteraceae;			
OC	Glucunobacter.			
OC	NCBI_Taxid=442;			
NP	11			
NP	SEQUENCE FROM N.A.			
RP	MEDLINE=92017653; PubMed=1833618;			
RX	Cleton-Jansen A.-M., Dekker S., Van de putte P., Goosen N.;			
RT	*A single amino acid substitution changes the substrate specificity			
RT	of quinoprotein glucose dehydrogenase in Glucunobacter oxydans.*;			
RT	Mol. Gen. Genet. 229:206-212(1991).			
RL	12			
RP	REVISION TO 213.			
RP	Goosen N.;			
RL	Submitted (FEB-1992) to the EMBL/Genbank/DBJ databases.			
CC	-1 CATALYTIC ACTIVITY: D-glucose + acceptor = D-glucono-1,5-lactone +			
CC	reduced acceptor.			
CC	-1 COFACTOR: POO.			
CC	-1 SUBCELLULAR LOCATION: OUTSIDE OF THE PERIPLASMIC MEMBRANE.			
CC	-1 MISCELLANEOUS: THE P1 FORM CAN OXIDIZE ONLY D-GLUCOSE, WHILE THE			
CC	P2 FORM CAN ALSO OXIDIZE MALTOSE. THE SEQUENCE OF P1 FORM IS SHOWN			
CC	HERE.			
CC	-1 SIMILARITY: BELONGS TO THE BACTERIAL POO DEHYDROGENASE FAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
CC	EMBL; X62710; CAA44594.1; ALT_SEQ.			
CC	PIR: S1716; OPKEX.			
DR	InterPro: IPR001479; Bac_POO.			
DR	InterPro: IPR002372; Bac_POO_repeat.			
DR	Pfam: pf01011; Bacterial_POO; 7.			
DR	PROSITE: PS00363; BACTERIAL_POO_1; 1.			
DR	PROSITE: PS00364; BACTERIAL_POO_2; 1.			
KW	Oxidoreductase; POO; Transmembrane; periplasmic; signal.			
FT	CHAIN	1	33	POTENTIAL.
FT	CHAIN	34	808	GLUCOSE DEHYDROGENASE [PYRROLOQUINOLINE-
FT	TRANSMEM	35	54	OUTINONE].
FT	TRANSMEM	59	76	POTENTIAL.
FT	TRANSMEM	94	108	POTENTIAL.
FT	TRANSMEM	123	138	POTENTIAL.

QY 237 PSTMEGLQSIINPNAIGERRVLTGVPCKTGMMPDAETGEEL-----WA 282
 DB 486 PS-----LVDFPKRDKDTTKPAVVGST--KSGGFYLDKRYVCKPLTKVLEQPIKADIPE 538
 QY 283 RDNNYQNMIESIDENGIVTNE-----DAIILEKD--VEYD-----VC 318
 DB 539 QYKSTQPRVEMEQIGNOLKESDMWGPATPFQOLMCRINFSKMRDGLYTPAGDVLSLF 598
 QY 319 PTFGLGHDPMSAALNPDGSIYF-----IPLN--NVCYDMMAVD 354
 DB 599 PGLSGNMWMSIAFDPTHRMYEFVNDRLGLMQLKQTPEDKIQANGKRYNTGMGAVP 658
 QY 355 QEFTSMQVYNTSVNKL--PPGKDMIGRIDAIDISTGRTLSVEERAAANS-----403
 DB 659 MKGTPYKNNKRRFMSALGIPCOKPPFGTMTAIDMKTRQVAMQVPLGTIODTGPAGIKMGL 718
 QY 404 -----PVLTGGGVLFNGCTDRYFRALSOETGETLQTRLATVASQAIISY--E 450
 DB 719 KAPIGMPTIGPMTQGLVFAATQDYLRATNNSNGKELMKARLPVGSQGTSPMSYSP 778
 QY 451 VDMQYVAIAGG 463
 DB 779 KTKQYVVSAGG 791
 RESULT 15
 QY 092078; STANDARD: PRT: 790 AA.
 ID QYIA_XANCI
 AC 092078;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Probable quinate dehydrogenase [pyrroloquinoline-quinone]
 DE (EC 1.1.99.25).
 GN QUMA.
 OS Xanthomonas campestris (pv. juglandis).
 CC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
 CC Xanthomonas.
 CC NCBI_TaxID=44291;
 OK [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=C5;
 RX MEDLINE=20063481; PubMed=10594704;
 RA Lee Y.-A., Lo Y.-C., Yu P.-P.,
 "A gene involved in quinate metabolism is specific to one DNA homology
 group of Xanthomonas campestris."
 J. Appl. Microbiol. 87:649-658(1999).
 RL
 CC -1- CATALYTIC ACTIVITY: Quinate + pyrroloquinoline-quinone = 5-
 dehydroquininate + reduced pyrroloquinoline-quinone.
 CC -1- COFACTOR: POQ (BY SIMILARITY).
 CC -1- PATHWAY: QUINIC ACID CATABOLIC PATHWAY; FIRST STEP. THIS PATHWAY
 ALLOWS GROWTH OF BACTERIA WITH QUINATE BY ITS CONVERSION TO
 PROTOCATECHUATE AND SUBSEQUENT METABOLISM BY THE BETA-KETOADIPIATE
 PATHWAY.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POQ DEHYDROGENASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation
 at the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 or send an email to license@sib-sib.ch).
 CC
 CC EMBL: AF109471; AAD38453.1;
 CC InterPro: IPR001479; Bac_POQ.
 CC InterPro: IPR002372; Bac_POQ_repeat.
 DR Pfam: PF01011; Bacterial_POQ_5.
 DR PROSITE: PS00363; BACTERIAL_POQ_1; FALSE_NEG.
 DR PROSITE: PS00364; BACTERIAL_POQ_2; FALSE_NEG.
 DR Oxidoreductase; POQ; Quinate metabolism; Transmembrane.
 KV TRANSMEM 22 42 POTENTIAL.

FT TRANSMEM 48 68 POTENTIAL.
 FT TRANSMEM 77 94 POTENTIAL.
 FT TRANSMEM 106 126 POTENTIAL.
 FT SEQUENCE 790 AA; 82896 MW; B75F29B52A49EE6F CRC64;
 SO
 Query Match 7.5%; Score 189.5; DB 1: Length 790;
 Best Local Similarity 21.2%; Pred. No. 8.6e-07;
 Matches 128; Conservative 57; Mismatches 204; Indels 215; Gaps 27;
 4 TSLMASAGAL-----ALLAAPFAQVTPVT-----DELLANPAGEWIS---43
 DB 238 TSSVWQRCRGIGYEDADALPAPVNPSPPLAAVYAQAGANCRRRLFTITIGRLIAYDA 297
 QY 44 -----YGNQENYRHSPLQITTEVNSQLDLVWARGMQPKQVOVTPLIHDGVYLAN 95
 DB 298 DTGAFCCGGGNS-----GVDLKAAGLGAAP-----DPTQLIS 330
 QY 96 --IVANGVIVAGST-----QYSPFGCVSGHDSATGEELMRNYFIPRAGEEDETWGD 148
 DB 331 PPLVACTVVGGRTRADDNVQYDMPGVVRG-----SMWSP--VRSAGLDP-----GNP 377
 QY 149 YE-----ARMMTGAMQIITYDPYTNLVHGSTAVGPASETORGCTGTYGYN--196
 DB 378 HDRQAPAGSSYVRSTPNWAMPASYDAAMTVF-----LPLGFSITLIGHERT 426
 QY 197 -----TRFAVRPPTGELVNRHOTLPRDNDQCTFEMMTVNVQVSTEMEGLQSIIN 248
 DB 427 ALDHRYSVASVIALDATYTGAEKRWYQTVHNDLMD-----FDLPMPQSL-----ID 470
 QY 249 PNAATGERRVLTGVPCKTGMMPDAETGEELW-ARDT-----NY--QNMIESI 294
 DB 471 FPNODGSHTPAVYIGIKAGQIYVLDNRATKPLTEVREPVKSGDIAHEQYAPTOPLISGM 530
 QY 295 DENGIVTNEDEAILKEIDVEYDVC-----PTFLGGRDMPSA 330
 DB 531 PGIQKHLTESDMWGPATPFQOLMCRIFAQKMRVYGLYTPAGTDVLSLSPGSGIAGNMWGL 590
 QY 331 ALNPDGSIYF-----IPLN--NVCYDMMAVDQFTSMQVYNTSVNKL--PPGKDMIGRIDAIDISTGRTLSVEERAAANS-----403
 DB 591 STDYVHDVVFANDMKRLGLMVOMIPADTRKAEAGGGAVENTGGAFLPKGTYYAVANKNF 650
 QY 368 VTKL-----PPGKDMIGRIDAIDISTGRTLSVEERAAANS-----YSPV-----405
 DB 651 LSAIGIPCOAPP-----YGTLSHIDKTSIMQVPAVGTVDOTGPGIKMHLPIPIGMPTL 706
 QY 406 ---LSTGGGVLFNGCT-DRYFRALSOETGETLQTRLATVASQAIISY--EVDGMQYVAI 459
 DB 707 GGTLSIOGGLIVTLAGQDYLRATNNSNGKELMKARLPVGSQGTSPMSYSP 778
 QY 460 AGG 463
 DB 767 SAGG 770

Search completed: May 24, 2002, 10:30:50
 Job time: 628 sec

Fri May 24 11:27:37 2002

walick-934-95.pep.rsp

Fri May 24 11:27:38 2002

walick-934-95.pep.rspt

Page 1

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2002, 10:19:27 ; Search time 93.8 seconds
(Without alignments)
870.508 Million cell updates/sec

Title: WALICK-934-95.PEP

Perfect score: 1 MKPTSLMASAGALALAAP.....GMOYVALAGGVSYSGSLNS 472

Sequence: BLOSUM62

Scoring table: Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: SP:archaea:19:***
2: SP:bacteria:***
3: SP:fungi:***
4: SP:human:***
5: SP:invertebrate:***
6: SP:mammal:***
7: SP:mhc:***
8: SP:organelle:***
9: SP:phage:***
10: SP:plant:***
11: SP:rodent:***
12: SP:virus:***
13: SP:vertebrate:***
14: SP:unclassified:***
15: SP:virus:***
16: SP:bacteriophage:***
17: SP:archaeal:***

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1380	54.5	608	2	093RE9
2	446.5	17.6	698	2	09KH03
3	435.5	17.2	695	2	09F902
4	430.5	17.0	629	2	09A048
5	428	16.9	601	2	09E1W8
6	427	16.9	601	16	092WY9
7	422.5	16.7	708	2	046444
8	421	16.6	742	2	053362
9	420.5	16.6	601	2	P71509
10	399.5	15.8	691	2	09A995
11	394.5	15.6	633	2	024759
12	386	15.3	599	2	091935
13	371	14.7	623	2	09AGW3
14	364.5	14.4	695	2	093460
15	359.5	14.2	573	2	059540
16	252.5	10.0	790	2	09X2S5

17	247	9.8	785	16	098KF6	098KF6 rhizobium l
18	237	9.4	803	16	091115	091115 pseudomonas
19	236.5	9.3	786	2	P95466	P95466 pantoea cit
20	217	8.6	777	16	092RB3	092RB3 rhizobium m
21	208.5	8.2	182	2	032699	032699 hyphomicrob
22	204.5	8.1	470	2	030326	030326 acetobacter
23	203.5	8.0	179	2	032697	032697 hyphomicrob
24	200.5	7.9	180	2	032696	032696 hyphomicrob
25	200.5	7.9	180	2	032700	032700 hyphomicrob
26	200.5	7.9	181	2	032615	032615 hyphomicrob
27	199.5	7.9	181	2	032692	032692 hyphomicrob
28	197.5	7.8	181	2	032703	032703 hyphomicrob
29	196.5	7.8	181	2	032621	032621 hyphomicrob
30	195.5	7.7	180	2	032706	032706 hyphomicrob
31	193.5	7.6	180	2	032701	032701 hyphomicrob
32	191.5	7.6	182	2	032612	032612 hyphomicrob
33	191.5	7.6	184	2	032693	032693 hyphomicrob
34	190.5	7.5	172	2	033894	033894 hyphomicrob
35	187	7.4	639	2	P77931	P77931 pseudomonas
36	186.5	7.4	172	2	033882	033882 methylosinu
37	186.5	7.4	179	2	032707	032707 hyphomicrob
38	185.5	7.3	180	2	032694	032694 hyphomicrob
39	185.5	7.3	180	2	032704	032704 hyphomicrob
40	185	7.3	171	2	033884	033884 methylosinu
41	183.5	7.3	179	2	032702	032702 hyphomicrob
42	182.5	7.2	184	2	032695	032695 hyphomicrob
43	181.5	7.2	185	2	0918K6	0918K6 uncultured
44	180.5	7.1	644	2	052551	052551 pseudomonas
45	176.5	7.0	172	2	093K58	093K58 methanotrop

ALIGNMENTS

RESULT 1
ID 093RE9 PRELIMINARY; PRT: 608 AA.
AC 01-DIC-2001 (TREMBLrel. 19, Created)
DT 01-DIC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DIC-2001 (TREMBLrel. 19, Last annotation update)
DE ALCOHOL DEHYDROGENASE.
GN ADH.
OS Pseudogluconobacter saccharotogenes.
OC Bacteria; Pseudogluconobacter.
OX NCBI_TaxID=133921;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IFO 14464;
RA Shidara T., Saito Y.;
RT "Alcohol dehydrogenase."
RT Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB046580; BAB62258.1; -
DR SEQUENCE 608 AA; 65101 MW; 0ACBCE97AE11BA570 CRC64;

Query Match 54.5%; Score 1380; DB 2; Length 608;
Best Local Similarity 46.7%; Pred. No. 1.2e-93;
Matches 272; Conservative 58; Mismatches 132; Indels 120; Gaps 5;
OY 10 SACALALAAPFAQ-----VFPTDELANDPAGEWISTYQONEN 50
DB 16 STALIASLGPFAFDHANAAPSKAQSAIENFQVTDLAGKPNWPTLRNGYOG 75
OY 51 YHSPILQTTTENVGOLQVARGMOPKQVPTPLHIDGYMYLAN----- 95
DB 76 WGISPLDQINKDVGLQVMSRTMEPSNCGAIAVNGVIFLANTINDVQALDKTSL 135
OY 96 -----
DB 136 IMEYRRRLPSASKFINSLSGAKRSTALRGDKYFVSMDFVALDARTGLAMETNRGOG 195
OY 96 -----IVANGVYAGSTQYSPGCVSGSHDSATGEELMRNYFTIPRAGEBDET 144

```

Db 134 KCDCDVVNRGVALYQKVFVGAFCGLAIDAANGKVKWEODTLVDRSKSYTTITGAPRVY 193
QY 99 NG-VIYAGSTCOYSPGCGVSGHDSATGEBLMRYFTIP-----RAGEEDETWQND 148
Db 194 NGKVIYINGGALEGVNG-VITAYDAETGSKQOMRYTYTGPDPARPFENEMAKAAATW--D 250
QY 149 YEAR-WMTG----ANGQITDPVNLVHYGSTANGPASESTORPGG-TLYGTNTFAVR 202
Db 251 PSCGKIYINGGGGTVMNTAFPELMLTIGTAGNAPSRKLRSFKGQDLYLAAY-VALN 309
QY 203 PRTGELVNRHQTLPDNDQCEFTEMVNTN-VDOPSTEMEGLOSINPMNATGRRRLT 260
Db 310 PRTGELVNRHQTLPDNDQCEFTEMVNTN-VDOPSTEMEGLOSINPMNATGRRRLT 260
QY 261 GVPCKYGTMOFDETEGFLMARDTNYONMIESIDENG-IYTVVEDAILKELDEYDVC 318
Db 352 NAF-KNGFEFVIDRTNKFISAKNFVDYVWAGSYDKNRPVETPDADTSGKPA---DVY 406
QY 319 PTEFGRRMPSALNPDGTYETPLNNVCYDMAVDOFTSMQVNTSN-----367
Db 407 PGPFGAHNMHSMFHPRLGLAFIPAOHYPL-TLADNKEM-----VHNKDSPEAHRYGNW 461
QY 368 ----VTKLPCKDMIGRIDALIDISTGRTLSYERAANYSPVLTGGCVLFNGCTDRYR 423
Db 462 LGMLVNAEPSPRSKPMGRLLAMPDLQAKVNRHDBGPNWGGLTATAGNLVFOGTADGRLY 521
QY 424 ALSQETGELTMOFRLATVASCQAISEVDGMQYVAIA-CGGVSYG 467
Db 522 AYHAATGKLMQAPFGSGVVAAPVYTLIDGRQYVAVAGVGYYG 566

RESULT 3
Q9P902 PRELIMINARY; PRT; 695 AA.
AC Q9P902;
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DE 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE ALCOHOL DEHYDROGENASE.
GN ADH.
OS Pseudomonas stutzeri (Pseudomonas perfectionarina).
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC NCBI_TaxID=316;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BC-2;
RA Chang C.-H., Herrick J.B., Okinaka R.T., Brainerd J.B.,
  Terwilliger T.C.;
RT "Identification and characterization of genes activated by 2-
  chloroethanol in Pseudomonas stutzeri BC-2."
RL Submitted (Aug-1999) to the EMBL/GenBank/DBJ databases.
EMBL: AF176640; AAC09249.1;
DR HSSP: Q9Z4J7; 1PIG.
DR InterPro: IPR002372; Bac_PQO_repeat.
DR InterPro: IPR000345; CytC_heme_bind.
DR Pfam: PF01011; Bacterial_PQO; 6.
DR Pfam: PF00034; Cytochrome_c; 1.
DR PROSITE: PS00190; CYTOCHROME_C; UNKOWN_1.
SQ SEQUENCE 695 AA; 75842 MW; D29A698A642E1A83 CRC64;

Query Match 17.2%; Score 435.5; DB 2; Length 695;
Best Local Similarity 25.0%; Pred. No. 8.2e-24;
Matches 140; Conservative 72; Mismatches 207; Indels 141; Gaps 18;
Db 15 ALAFAFAQVTPVTEBL--ANPAGEWISGQENYRHSPLTOITTEVGOLOLWNA 72
QY 16 ALIYA-AGAAVAVDEAIAIRASQDSSEWLSHGRTVAEORRSPKQIDAGNVKTLANAY 74
Db 73 RGMQPGK-VQVTPPLIHGVWYLA-----94

```

```

Db 134 KCDCDVVNRGVALYQKVFVGAFCGLAIDAANGKVKWEODTLVDRSKSYTTITGAPRVY 193
QY 99 NG-VIYAGSTCOYSPGCGVSGHDSATGEBLMRYFTIP-----RAGEEDETWQND 148
Db 194 NGKVIYINGGALEGVNG-VITAYDAETGSKQOMRYTYTGPDPARPFENEMAKAAATW--D 250
QY 149 YEAR-WMTG----ANGQITDPVNLVHYGSTANGPASESTORPGG-TLYGTNTFAVR 202
Db 251 PSCGKIYINGGGGTVMNTAFPELMLTIGTAGNAPSRKLRSFKGQDLYLAAY-VALN 309
QY 203 PRTGELVNRHQTLPDNDQCEFTEMVNTN-VDOPSTEMEGLOSINPMNATGRRRLT 260
Db 310 PRTGELVNRHQTLPDNDQCEFTEMVNTN-VDOPSTEMEGLOSINPMNATGRRRLT 260
QY 261 GVPCKYGTMOFDETEGFLMARDTNYONMIESIDENG-IYTVVEDAILKELDEYDVC 318
Db 352 NAF-KNGFEFVIDRTNKFISAKNFVDYVWAGSYDKNRPVETPDADTSGKPA---DVY 406
QY 319 PTEFGRRMPSALNPDGTYETPLNNVCYDMAVDOFTSMQVNTSN-----367
Db 407 PGPFGAHNMHSMFHPRLGLAFIPAOHYPL-TLADNKEM-----VHNKDSPEAHRYGNW 461
QY 368 ----VTKLPCKDMIGRIDALIDISTGRTLSYERAANYSPVLTGGCVLFNGCTDRYR 423
Db 462 LGMLVNAEPSPRSKPMGRLLAMPDLQAKVNRHDBGPNWGGLTATAGNLVFOGTADGRLY 521
QY 424 ALSQETGELTMOFRLATVASCQAISEVDGMQYVAIA-CGGVSYG 467
Db 522 AYHAATGKLMQAPFGSGVVAAPVYTLIDGRQYVAVAGVGYYG 566

RESULT 2
Q9KH03 PRELIMINARY; PRT; 698 AA.
AC Q9KH03;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DE 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE TERAHYDROFURFYL ALCOHOL DEHYDROGENASE.
OS Alcaligenes eutrophus (Ralstonia eutropha).
OC Bacteria; Proteobacteria; beta subdivision; Ralstonia group;
OC NCBI_TaxID=510;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-21123557; PubMed=11222593;
RA Zarnit G., Schrader T., Andreesen J.R.;
RT "Catalytic and Molecular Properties of the Quinohemoprotein
  Tetrahydrofurfuryl Alcohol Dehydrogenase from Ralstonia eutropha
  Strain Bo."
RL J. Bacteriol. 183:1954-1960(2001).
EMBL: AF277373; AF86335.1;
DR HSSP: Q9Z4J7; 1PIG.
DR InterPro: IPR002372; Bac_PQO_repeat.
DR InterPro: IPR000345; CytC_heme_bind.
DR Pfam: PF01011; Bacterial_PQO; 6.
DR Pfam: PF00034; Cytochrome_c; 1.
DR PROSITE: PS00190; CYTOCHROME_C; UNKOWN_1.
SQ SEQUENCE 698 AA; 75471 MW; 9C87F180DBD12FB2 CRC64;

Query Match 17.6%; Score 446.5; DB 2; Length 698;
Best Local Similarity 25.5%; Pred. No. 1.3e-24;
Matches 149; Conservative 74; Mismatches 203; Indels 159; Gaps 21;
Db 10 SACALALLAFAFA--QVTPVTEBLANPPAG--EWISGQENYRHSPLTOITTEVNG 65
QY 14 AASVALPAPAFGANAAARVDGAIRANEAGTPNPNSTGLDAEYERFSEKLEQVAVANVR 73
Db 66 QQLVWAKMPGK-VQVTPPLIHGVWYLA-----93
QY 74 NGLAMSTDELESTRGVETPLVVDGVVYSAFMSVVAHALDARTGLMTYDPOVPODAY 133
Db 94 -----ANIVA 98

```

```

Dd      11  TGVSVAAALAAALPLAGPS-----ALANDLEVELSKSDGNVWPKQNTVDSDNTSKUKOJNAE    66
Dd      63  NVGOLQIYV--ARQMOPGKVOYVPLIHGVNVL-----                               93
Oy      11  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Dd      67  NKNKLNKVSQMSQSTGLNGH--EGAPLVVDGDMTVHTSEPNNTFALGDDEKTLIMODPKQ    125
Oy      94  -----ANIV-----NGVIVA    104
Dd      126  NPAARSACCDLVNGLAWMGDGKTPSLITLTLDDGHVVALNAOQETWVKIENSDIRV    185
Oy      105  GSTRCOXSPF-----GC-----FVSGHDSATGEBELMRNYFLPRAGEGD-----    142
Dd      186  GSTLTTPYVYMDKVIIGSSAGELGVAGYTLAVDTGQCKRAY--AAGPSDLLAK    242
Oy      143  -----ETWNGVDYFARMTCG--NGOILYDVPINLVHYSIAGVASPT    183
Dd      243  DENIHNAHYGOKGLSTWEED--ANKIGGTGNMGWYAVDGPNTLIFYGNDPAWNET    299
Oy      184  ORGPRGGLTGVITREPAVRPDPTGELVNRHOTLRDDMDDECFEMWYKINVDVOSTEMBG    243
Oy      300  MR--PGDNKR--TWITIRARVDVGEKKFYCTPRDEXDVAQVNVNVL-----STOKD--    348
Dd      244  LOSINPMAATGERVULGVACVCTGTMMQFADETGELMAR--DPNYNMIESIDENGIV    300
Oy      349  -----RSKREKKLLTHDRNGIYUILLDRNTODLISAKHIDT--VNRKTYDLKSGL    398
Dd      301  TVNEDALLKEID--VEYVOVCTPLGSDRMPMSALNPDGSIFFIPLNNVCYDMMANDQETS    359
Oy      399  PVADPEYGRTRDHLAKDICSFMAGYNGHDSYDERKLYTGMGHNICDMWEPMLPYRA    456
Dd      360  MDVYNTSNVTKLPRGK-----DMIGRIDAIDISGRTLVSEVERAANSPVLSGGCVL    413
Oy      459  GQFVQATITNMPGPKGDRONABGLGAIKAYALITGRKKWEMKREFWNVGCTLATAGNV    516
Dd      414  FNGGTGRFPRAISOETGETTLMOQLATVAISGALISYVDWQMOVAIVAGGSVSGS    469
Oy      519  FYGLDGFIAKHSPTGELIMAKPLGSAIGVPIYTHKGIOYAI-----YGVG    569

```

	RESULT	5			
D	OGEYW8		PRELIMINARY:	PRT:	601 AA.
C	OGEYW8:				
T	01-MAR-2001 (TIREMBrel, 16,				
T	01-MAR-2001 (TIREMBrel, 16,				
T	01-DEC-2001 (TIREMBrel, 19,				
E	METHANOL DEHYDROGENASE LARGE SUBUNIT-LIKE PROTEIN.				
N	KMAAF				
N	Rhizobium meliloti (Sinorhizobium meliloti).				
S	Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;				
C	Rhizobiaceae; Sinorhizobium.				
C	NCBI_TaxID=382;				
CX	[1]				
RRN	SEQUENCE FROM N.A.				
RRP	SPRAIRV-RM1021;				
RC	Fenner B.J., Tiwari R.P., Dilworth M.J.;				
RT	"Regulation of Cl assimilation in Sinorhizobium meliloti.";				
RL	Submitted (SEP-2000) to the EMBL/Genbank/DBSJ databases.				
RL	EMBL: AF309488; AAC31643.1; "				
DR	HSSP: P38539; AAAH.				
DR	InterPro: IPR002372; Bac_PQQ-repeat.				
DR	Pfam: PF01011; Bacterial_POQ_7.				
DR	SEQUENCE 601 AA; 65686 MW; 09F743726B2E2A83 CRC64;				
StQ					

	Query Match	16.9%;	Score 428;	DB 2;	Length 601;
	Best local Similarity	24.9%;	Pred. No. 2,4e-73;		
	Matches 149; Conservative	72;	Mismatches 202;	Indels 176;	Gaps
QY	14 LALLAAPFAAQVTPTVDL--LANPAGEMLISGCGDENTRRSEPLQTITTEWVGOLDIWF 71				
db	8 LAIMSIGGAQAVAFANDDELQKLTIDP-NQMALIQGTIANLRYSKLQDIKNKDVMKGLQAW 66				

	22
--	----

Db 8 LAIMVSGGAQAFAFANDELQKILDDP--NQAVIDGTQYANLRKSLDQINDKNVSKLOVAM 66
QY 72 A-----RCMQ-----PGVQVTPRLIH----- 88
Db 67 TFGTGVLRGHEGSPLYVIDDLKATYTPPTPTVTVALDLSKGOIWMKEPKOPNVIPMCC 12
QY 89 -----GVMLAN-----IYANGVVGST----- 10
Db 127 DIVNNGVATADKQKLFELHQADITVVALDAKTGKIVMSVANKCATGGEINATITVPYKIL 18
QY 108 --CQSPFGC--FVSGHDSATGELMRNFEI-----PRAGEGDETW 14
Db 187 VGISGGEVGVGHVATVASMADGKVLWRGISMKQPSDSDLIDPEKTTILGKPKVCKDSGLTW 24
QY 146 GNDYFARMNG-----AMQITDPATNLVHGSTAVGASPEFOTKGTGGLYGNTRFAVR 20
Db 247 GSD---QMKIGGGTIMGWSTYDPEENLYVVGSTGNPSTMNPTQR--PGDKRM-SMTIFARD 30
QY 203 PDGTGEIWRHQTLPDRNMQDECTFFEMATNTVDVQPSFEMGLDSINPNAATGERRVLNGV 26
Db 301 VDTMAKMLQMPHPDEMIDYGVNMLITEQIDGK-----DRKLTLHF 34
QY 263 PKGTGTMMQGVDAETGEL-----WADNTNYQMI-----SIDENITV 30
Db 345 D-FNGGTYMDRYTGELLVAERKYDPTVMNATEVWMPKSDKXGRQVAAQVSTQDN--- 40
QY 302 VNEDAILKELDEVDCPTFLGGRDMPSALINDSGITFIPLANVCIDMAVDQEFSTMD 36
Db 401 --EDNTTT-----GVCPAALGTQDQAPAYSKTELEFVPLPHNHCMDYEFRRYSTAQ 42
QY 362 VYNTSVNTELPKQKDM---IGRIDALDISTGRTLSVERAANYSPLSTGGGVLTNGST 48
Db 453 PVYGATLSWYP-KDSHGKGNFIANDMNEGKIKWSLPEPSVSGALATAGDVYETGL 51
QY 419 DRYFALSOEGLFELMQLRATLVAASQAISYEVDDMOYVATAG-----CVYSGSGLNS 47
Db 512 EGYLAADAAGAKELRYFKTPSGVLIGNVMTYARBSKQVAVLSGVGMAGIGLAGLTN 57
RESULT 7
Q46444 PRELIMINARY; PRT; 708 AA.
ID Q46444
DT 01-NOV-1998 (TEMBLrel. 08, Created)
DT 01-NOV-1998 (TEMBLrel. 08, Last annotation update)
DT 01-DEC-2001 (TEMBLrel. 19, Last annotation update)
DE QTNQINHAOPROTEIN ETHANOL DEHYDROGENASE TYPE I PRECURSOR
DE (EC 1.1.99.-) (OH-EDH1),
OS OHEDH,
GN Comamonas testosteroni (Pseudomonas testosteroni),
OS Bacteria; Proteobacteria; beta subdivision; Comamonadaceae; Comamonas.
OC NCBI_TaxID=285;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15667;
RA MEDLINE=96184549; Pubmed=8654419;
RA Stoorvogel J., Kraayveld D.E., van Sluis C.A., Jongejan J.A.,
RT dehydrogenase of Comamonas testosteroni.";
RN Eurt. J. Biochem. 235:690-698(1996).
RP SEQUENCE OF 32-54 AND 477-490.
RC STRAIN=ATCC 15667;

RA De Jong G. J. 1980; PMID=6701151;
 RT Duine J. A. J. H., Geerlof A., Stoorvogel J., De Vries S.,
 RT "Quinohemoprotein ethanol dehydrogenase from *Comamonas testosteroni*.
 RT Purification, characterization, and reconstitution of the apoenzyme
 RT with pyroloquinoline quinone analogues.";
 RL Eur. J. Biochem. 230:899-905(1995).
 RN [3]
 RP CHARACTERIZATION.

RA MEDLINE-86242113; PubMed-3521592;
 RX Groen B.W., van Kleef M.A., Duine J.A.;
 RT "Quinolone protein alcohol dehydrogenase apoenzyme from pseudomonas
 testosteroni".
 RL Biochem. J. 234:611-615(1986).
 RN [4]
 RP CRYSTALLIZATION;
 RA MEDLINE-21336088; PubMed-11679760;
 RX Oubrie A., Hultinga E.G., Rozeboom H.J., Kalk K.H., de Jong G.A.H.,
 RA Duine J.A., Dijkstra B.W.;
 RT "Crystallization of quinohemoprotein alcohol dehydrogenase from
 Comamonas testosteroni: crystals with unique optical properties";
 RL Acta Crystallogr. D 57:1732-1734(2001).
 CC -1- FUNCTION: CATALYZES THE DYE-LINKED OXIDATION OF PRIMARY ALCOHOLS
 TO THE CORRESPONDING ALDEHYDES AND THE (SUBSEQUENT) OXIDATION OF
 THE ALDEHYDES TO CARBOXYLIC ACIDS.
 CC -1- COFACTOR: PQQ, HEME, AND CALCIUM.
 CC -1- SUBUNIT: MONOMER.
 CC -1- SUBCELLULAR LOCATION: PERIPLASMIC (POTENTIAL).
 CC -1- INDUCTION: BY ETHANOL AND BUTANOL.
 CC -1- SIMILARITY: TO ALCOHOL DEHYDROGENASE (ACCEPTOR) PROTEINS.
 CC EMBL: X81880; CA57464.1; .
 DR HSP: Q9Z4J7; IFLG.
 DR InterPro: IPR002372; Bac_PQQ_repeat.
 DR InterPro: IPR003088; Cyt.C1.
 DR InterPro: IPR002329; Cyt.C1.
 DR Pfam: PF01011; Bacterial_PQQ_6.
 DR Pfam: PF00034; Cytochrome_c_1.
 DR PRINTS: PR00605; CYTOCHROME.C1.
 DR Signal: PQQ; Heme; Calcium; Oxidoreductase; Periplasmic.
 KW SIGNAL: 1 31
 FT CHAIN: 32 708 QUINOHEMOPROTEIN ETHANOL DEHYDROGENASE
 FT TYPE 1.
 FT BINDING: 635 635 HEME (COVALENT) (BY SIMILARITY).
 FT BINDING: 638 638 HEME (COVALENT) (BY SIMILARITY).
 FT METAL: 639 639 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
 FT SEQUENCE: 708 AA; 76822 MW; 99AB5BDD6ACCA3 CRC64;

Query Match 16.7%; Score 422.5; DB 2; Length 708;
 Best Local Similarity 23.5%; Pred. No. 7.7e-23;
 Matches 139; Conservative 79; Mismatches 222; Indels 151; Gaps 18;
 3 PSLIMASAGALLAFAAQTPTVDEL--DELANP-AGENISYGOQEN 50
 12 PERWVLLAACIG--SAAAFAGTGPAAQAAAVORVDGFIRANARPPDPTIGVYAE 69
 51 YHSPLEQTITENYGOLOLVARQMOPK-VQVPLIHGVMTLA----- 94
 70 TRYSRLDQINANAKDGLAMSYNLSTRGVEATPVVDGIMVYSASVVAHIDRTGN 129
 95 -----NIVANG----- 100
 130 RIWYDPOIDRSTGFGCCDVGVALMKQYVVGANORLIALDAATGKEVMEQNTFE 189
 101 -----VIVAGSTCOYSPGCFVSGHDSATGEELMRNYETP----- 135
 190 GQKSLITIGARVPEKGVIIIGKRGAEYVGR-YITAYDAETGRKWRMSVPGDPSKRF 248
 136 --RAGEGDETDGNDYEAMMTGA-----WQITVDPVPTNLVHYSYAVGPASSETORTP 188
 249 EDESMKRAARTW--DPSGKMWEGAGGGGTWDSMTDALNMYGTGSGPMSHKYRSPK 306
 189 GGLTGTNTFAVRPDTEIWRHQTLPDNDQECTEFEMAVTVVDVQPSRMEGLQSTN 248
 307 GGDNLVLAISALDPDGTGKMYHETPDNDNDYSTPMLADIKT----- 353
 249 PNAATGERVLTGVPCCTGTMMQFDETEGFLIARDTVYQNMIESIDENG-IYTVNEDAI 307
 354 --AKPRKVIHAP-KNGFFVLDRTNGKFKISAKNFVYVNNASGYDRHGKIGI--AA 406
 308 LKEDLVERDVCPTFLGGRDWPSSALNDGSIYFIPLNVCYMMAVDQ-EFT----- 358

DB 407 ARDSKPODAVPGPYGAAHNMHSPNFDPOTGLVLEPAQNVPNLMDKKMEFNQAPGKPO 466
 359 SMDVYNTSNVTKL-PQCKMDIGRLDAIDISTGRITLVSERAAANYSPVLTGGVYFENG 417
 467 SGTGWNIAKFFNAEPKSPKPRGLAMDPAQAAAMSVHVSFPMNGGTLITAGNVYQGT 526
 418 TDRYRALSOETGETIMQTRLATVAGSAISYVDGKQYVLA-GGGSYSG 467
 527 AGRLVYHATATGKIMPAPTGTGVAAFPSTYVWDGQYVSANVAGGYG 577
 RESULT 8
 053362 PRELIMINARY; PRT; 742 AA.
 ID Q53362
 AC Q53362; Q44159;
 DT 01-NOV-1996 (TREMBLER, 01, Created)
 DT 01-NOV-1996 (TREMBLER, 01, Last sequence update)
 DT 01-OCT-2001 (TREMBLER, 18, Last annotation update)
 DE DEHYDROGENASE SUBUNIT OF ALCOHOL DEHYDROGENASE.
 OS Acetobacter pasteurianus (Acetobacter turbidans).
 OC Bacteria; Proteobacteria; alpha subdivision; Acetobacteraceae;
 OC Acetobacter.
 OX NCBI_Taxid=438;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN:NC11380;
 RC MEDLINE-94042848; PubMed-8226628;
 RX Takemura H., Kondo K., Horinouchi S., Beppu T.;
 RA "Induction by ethanol of alcohol dehydrogenase activity in Acetobacter
 pasteurianus".
 RT J. Bacteriol. 175:6857-6866(1993).
 RL J. Bacteriol. 175:6857-6866(1993).
 DR EMBL: D13893; BAA4052.1; .
 DR HSP: Q9Z4J7; IFLG.
 DR InterPro: IPR001479; Bac_PQQ.
 DR InterPro: IPR002372; Bac_PQQ_repeat.
 DR Pfam: PF01011; Bacterial_PQQ_6.
 DR Pfam: PF00363; BACTERIAL_PQQ_1; 1.
 DR PROSITE: PS00364; BACTERIAL_PQQ_2; 1.
 DR PROSITE: PS00364; BACTERIAL_PQQ_2; 1.
 DR SEQUENCE: 742 AA; 81680 MW; 031C4E9B6B13EC98 CRC64;

Query Match 16.6%; Score 421; DB 2; Length 742;
 Best Local Similarity 24.7%; Pred. No. 1.1e-22;
 Matches 150; Conservative 72; Mismatches 205; Indels 180; Gaps 20;
 11 AGALLAFAAFAQVPTVDEL--ANPAGEMISYGOQENYHSPLEQTITENYGOLO 68
 23 AALPYAAVPAADAGGQNTGEALIHADHPENWLSGTYSERYSPLDQINRSNGDLK 82
 69 LVARQMOPKQVQ-VPLIHGVMTLA-----ANI----- 96
 83 LAMWYTTLDTRNGQZATPLVVDGIMVATNMWSKMEALDAATGRLIMQYDPKVPINADKC 142
 97 ----- 96
 143 CQTVNRKAGYNGKGVFWGTEDGRLVAADAKTGKKEVNTIPADASLKORSYTVGAVR 202
 97 VANGVIVAGSTCOYSPGCF--FVSGHDSATGEELMRNYETPRAAGEGD----- 142
 203 VAKGLVLIN--GSEFPAARGVSAFDAETGKLMKRYTVYVNNKNEPDAVANDYLMASKA 260
 143 -ETWMDYEARMT-----GANGQITVDPVPTNLVHYSYAVGPASSETORTP 193
 261 YKTWGP--KGAVVRGOGGQVWDLSLVDPVSLLY--LAVG-----NGSPWNNKYRSE 309
 194 --GTN-----TRFAVRPDTEIWRHQTLPDNDQECTEFEMAVTVVDVQPSRMEGLQSTN 247
 310 GIGSNLEFLGSIYALKRETEGEYVHFQATPMDQWYTSVQOIMTLDMFY----- 357
 248 NPNAATGE--RVLTVGPCCTGTMMQFDETEGFLIARDTVYQNMIESIDENG-NCIYTVNED 305
 358 -----NCEMRHVIMHAP-KNGFFVLDRTNGKFKISAKNFVYVNNASGYDRHGKIGI--AA 411

```

QY 306 ALTKELVEYDVCPFLGDRMPGSAALNPDGTYTFLPANNVCY-----DMA 352
Db 412 GLYTLTGKFWYGIPIGPIGANNFMKMAVSRTHLVYPAHQIFPGKYNQGVGFKPHDPAMN 471
QY 353 VDOETSMVYNTSNVTKLPKPGDMIGRIDAIDISTGRTIMSEVFAAANYSPLYSTGGV 412
Db 472 VGLDMTKKGLDPE-ARTAYIKDLHGLMLANDPVKMETVKIDHKGWNGCVLATGDL 530
QY 413 LFNCGDTRFYRALSOETGETIMOTRLATVAGSQAISTEYDQMYVA-----IAG 462
Db 531 LFGDLANGEFHAYDATNGSDLYKFPDAGSIIASPMYTSVYNCKQYVAEVMGCIPLSMG 590
QY 463 GVSYSGL 469
Db 591 GVGRTSG 597

```

```

RESULT 9
P71509
ID P71509 PRELIMINARY; PRT: 601 AA.
AC P71509;
DT 01-FEB-1997 (TREMBLrel. 02, Created)
DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)
DE METANOL DEHYDROGENASE LARGE SUBUNIT HOMOLOG.
GN MXARF.
OS Methylobacterium extorquens.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Methylobacterium group; Methylobacterium.
OX NCBI_TaxID=408;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AMI;
RA MEDLINE=97312011; PubMed=9168622;
RA Christoserdova L., Lidstrom M.E.;
RT "Molecular and mutational analysis of a DNA region separating two
RT methyloctrophy gene clusters in Methylobacterium extorquens AM1.";
RL Microbiology 143:1729-1736(1997).
DR EMBL: U72662; AAB58890.1;
DR HSSP: P38539; AAAH.
DR InterPro: IPR002372; Bac_PQQ_repeat.
DR Pfam: PF01011; Bacterial_PQQ.
SQ SEQUENCE 601 AA; 64952 MW; 68E45C7059CB8239 CRC64;

```

Query Match
Best Local Similarity 16.6%; Score 420.5; DB 2: Length 601;
Matches 147; Conservative 90; Mismatches 203; Indels 157; Gaps 24;

```

QY 1 MKPSTLWASAGALALLAARPAQVPTVDELIANPAGEMISYQONENRHSPLQIT 60
Db 1 KRAVHLALAGGLAA--ASPLANESVLG--VANP--AEGLVQVVDYANTRYSKLDQIN 54
QY 61 TENVGQQLQVW--ARGMOPGKVQVPLIDHGYWL--ANIV-----97
Db 55 ASNVKNLQVAMWTESTGLRGH--EGSPLVYGNIMYHTPPNIVAYALDLDOGAKIWKYEP 113
QY 98 -----ANGVIV-----106
Db 114 KQPSVTPWCCDTVRKGLAYADGAILLHQAQDTLLVSLDAKSGKVMNSVANQDPSGETN 173
QY 107 TCQYSP-----FG--CFVSGHDSATGEELMRNYFI-----134
Db 174 TATVLPVKDKVLYIGISGFGVQCHYATYALDKSGKKWRRYSIGPDQILVPEKTSIG 233
QY 135 -PRAEGEGETMGNDYERAMWTG--AMGQITVDPVNLVHYGSTAVGASSETQCTGCG 190
Db 234 KPIADKSSIKTWEED--QMKTGCGGCTMGWFSYDPKIDLMYISG--NFTWNPQRQGD 288
QY 191 TLYGTNTRAVAPDPTGELIWRHOFLLPRDNDQCTEAMATVNDVQPSSTEMGLQSIIPN 250
Db 289 NKM--SMTIAMNPDITGAKAKVYIQMTPHDEWDFGINEMILTD-----QKTDG-----334

```

```

QY 251 AATGERVLTGVCCKTGTMMQFDAQETGEFLARDTN-YQNNIESID-ENGIVTNEADLI 308
Db 335 ---KDRPLLHPD-RNGFGTTLDRATGEVLAKEFDPVYNNATKVDLDKSKTYGRPLV 390
QY 309 KELDVEID-----VCPFLGDRMPGSAALNPDGTYTFLPANNVCYDMAAVDEFTSM 360
Db 391 SKYSTQONEDVNSKQICPAALGTQDQPAFSPKTELFYPTNHVCMDYEPFRVLYTPG 450
QY 361 DVYNTSNVTKLP-PGK-DMIGRIDAIDISTGRTIMSEVFAAANYSPLYSTGGVLFNGCT 418
Db 451 QPYVATLISMYPAPSGHSGKGNFLAMDNLQCKIKWSNPEQFSAMGALATSGDVLFTGL 510
QY 419 DRYFALSOETGETIMOTRLATVAGSQAISTEYDQMYVAIAG-----GVSYSGL 470
Db 511 EGFLEKAVDSKTGKELKFKTFPSGIIQVMTYEHKQKHVAVLGSGWAGIGLAAGL 567

```

```

RESULT 10
O9AF95
ID O9AF95 PRELIMINARY; PRT: 691 AA.
AC O9AF95;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DE 1-BUTANOL DEHYDROGENASE BDH.
OS Pseudomonas butanovora.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OX NCBI_TaxID=86174;
RN [1]
RP SEQUENCE FROM N.A.
RA Vanaul A.S., Arp D.J., Sayavedra-Soto L.A.;
RT "Characterization of the expression of two distinct alcohol
RT dehydrogenases involved in butane metabolism in Pseudomonas
RL butanovora.";
DR Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF355796; AAK27220.2;
SQ SEQUENCE 691 AA; 75070 MW; 4FCFED20CA1AE64 CRC64;

```

Query Match
Best Local Similarity 15.8%; Score 399.5; DB 2: Length 691;
Matches 137; Conservative 74; Mismatches 191; Indels 147; Gaps 19;

```

QY 39 GEMSTYQONDENRHSPLTQITTEENVGOLVMARGMOPGK-VQVTPLIHDGYWL-----93
Db 41 GEMRTHGYDAGTRYSPLAQITPDNAKELGLVMSIDLESSRGVATPIVVDGYVYTAEP 100
QY 94 -----ANIVANGVIV-AGSTQYSPGCGFVSGHDS 122
Db 101 SVYHALDVRGKRLMTYDEVRERKGNACCDVYNRGVAHEGKVPFVSGIDGLVA-IDA 159
QY 123 ATGGEIWM-RNYFI-----PRA--GEQGETMGNDYERAMWTGAM-----GO 160
Db 160 RTGKRWERNTLIDDKPYTTIGAPRYKGVVINGNGAEFGVGYTAYDPTAASRGV 219
QY 161 I-----TYDPTNLVHYGSTAVG-----178
Db 220 VEPQDPSLPFEDASMEAAKTMWDPAQVIGSGRRHGVELQILRKAGFCTSAAPAPS 279
QY 179 PASETQRTPGGTLTGNTREAVRPDTEGLVNRHQLPLPDNDQCTEAMATVNDV--Q 236
Db 280 PWSHRKSPAGGDMNLTAISYALRPDTEGLVYHYYQOTPADNDYVSTODLILADIELGK 339
QY 237 PSTHEGLQSTINPANAAGERRVLTVGPCYKTMQCFDAQETGEFLMARPTQNNIESIDE 296
Db 340 P-----RKVILHAP-KNGFFVYIDRTDGFISAQNFYVVMNATGYDE 380
QY 297 NGIYTVN-EDAILKELVEYDVCPFLGDRMPGSAALNPDGTYTFLPANNVCYDMAVAD 355
Db 381 NGRPIENPEGAMFGHLSMR--PAFSARTNWSMSYSPQGLAVFPQAQNIPL-VLQEDK 436
QY 356 EFTSMVYNTSN-----VTKLPKPGDMIGRIDAIDISTGRTIMSEVFAA 400

```



```
Db 437 NMS-----YNOAGQAGMAGIGNMLGNVPRPBPQSPGRILANDPVOQEVNRKREHVEP 492
OY 401 NYSVLTSTGGVLFNGSTDRFALSOETGETLWOTRLATVASQALSYEDVQGVYVLA 460
Db 493 WNGSLTVAGNVFQGTADARLLAFAPARDGKBLMSAPMGTVIAAPVTEYVDKQVISA 552
OY 461 -GGGVSYGS 468
Db 553 VGVGVGVYGN 561

RESULT 11
ID 024759 PRELIMINARY: PRT; 633 AA.
AC 024759;
DT 01-JAN-1998 (TREMBLER, 05, last sequence update)
DT 01-JAN-1998 (TREMBLER, 19, last annotation update)
DT 01-DEC-2001 (TREMBLER, 19, last annotation update)
DE METHANOL DEHYDROGENASE ALPHA SUBUNIT PRECURSOR (Ec 1.1.99.8).
GN MXAF.
OS Bacterium methylotrophicum.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Bacterium methylotrophicum.
OC Bacterium methylotrophicum.
OC NCBI_TaxID=84;
OX 11
RN SEQUENCE FROM N.A.
RC STRAIN=GM2;
RA MEDLINE=97457202; PubMed=9311140;
RA Tanaka Y., Yoshida T., Matsumoto K., Izumi Y., Mitsunaga T.;
RT "Cloning and analysis of methanol oxidation genes in the methylotroph
RT Bacterium methylotrophicum GM2."
RT FEMS Microbiol. Lett. 154:397-401(1997).
DR EMBL: AB004097; BAA23272.1;
DR HSSP: P38539; 4AAH.
DR InterPro: IPR001479; Bac_PDO_repeat.
DR Pfam: PF01011; Bacterial_PDO_7.
DR PROSITE: PS00363; BACTERIAL_PDO_1; 1.
DR PROSITE: PS00364; BACTERIAL_PDO_2; 1.
DR Signal: oxidoreductase.
KW SIGNAL.
FT SIGNAL 34
FT CHAIN 35 633 POTENTIAL.
SQ SEQUENCE 633 AA; 69853 MW; B47A3A279E2C1B CRC64;
```

```
Query Match 15.6%; Score 394.5; DB 2; Length 633;
Best local similarity 23.6%; Pred. No. 7.6e-21;
Matches 140; Conservative 88; Mismatches 213; Indels 151; Gaps 21;

OY 6 LMSAGALALLAFAFAVPTVDELALNPAGE-WISYQONQENYRHSPLTOITTEV 64
Db 12 LMSAGALALLAFAFAVPTVDELALNPAGE-WISYQONQENYRHSPLTOITTEV 71
OY 65 GOLULVNA----- 72
Db 72 KOLKAMFSTGELHGHGAPLVIGDMYVHSSFPKTFALNLPDGHILMOSHBPQDPA 131
OY 73 -----RGMO--PGKVOTPLI--HDGVYTLANT-----VANGYIVAGST 107
Db 132 ARSVACDLVNRGLAYVPGDKPVALYIKTOLDGLHVALNATGEEFMKVEGDIKVGOT 191
OY 108 COVSPGCF-----VSGHDSA-----TGEIMRNRYFI----- 134
Db 192 LTPAPVVDLAIYVSGSAGELGVGRHAYAVNKGEGAMRYVATGDEITGLADDFNSN 251
OY 135 PRAGEE--GDETMNDYEARMWGA--WGQITVPTNLVHGSTAVGAPASETOGTGTPG 189
Db 252 PHTGQGLGATWEGD--AMKIGGTWGWAYDPOANLITYGSGNPANWETAR--PG 306
OY 190 GLTGTNTRFAPVPTGELVWHRQTLPRDNDQCTEFEMVTVNDVOPSTEMGLQSLNP 249
Db 307 DNKW-TWTITRADDTGKMKFGYKTPHDEMDPAGVYIMLSE-----QTDKE----- 353

RESULT 12
ID 09L935 PRELIMINARY: PRT; 599 AA.
AC 09L935;
DT 01-OCT-2000 (TREMBLER, 15, last sequence update)
DT 01-OCT-2000 (TREMBLER, 15, last sequence update)
DT 01-DEC-2001 (TREMBLER, 19, last annotation update)
DE METHANOL DEHYDROGENASE LARGE SUBUNIT.
GN MXAF.
OS Bacterium methylotrophicum sp. (strain SSI / DSM 11726).
OC Bacteria; Proteobacteria; beta subdivision; Methylophilus group;
OC Bacterium methylotrophicum.
OC Bacterium methylotrophicum.
OC NCBI_TaxID=81683;
OX 11
RN SEQUENCE FROM N.A.
RC STRAIN=SSI;
RA Kim Y.M.;
RT "Cloning and nucleotide sequence of mxar gene of Methylophilus sp.
RT submitted (SEP-1998) to the EMBL/Genbank/DBJ databases.
DR EMBL: AF184915; AAD56237.2;
DR HSSP: P38539; 4AAH.
DR InterPro: IPR001479; Bac_PDO_repeat.
DR Pfam: PF01011; Bacterial_PDO_7.
DR PROSITE: PS00364; BACTERIAL_PDO_2; 1.
DR PROSITE: PS00364; BACTERIAL_PDO_2; 1.
SQ SEQUENCE 599 AA; 65133 MW; DBF6F4B5D871BC91 CRC64;
```


Db 298 W-TMTIARDVDYGAKKWYQKTPHDEMDPAVQNMILTDQAVNGKTY----- 344

OY 253 TGBRRLGVCKTGMQDPAETGEFLMARDTN-YQNMIESIDENGITVNEADILTEL 311

Db 345 -----FLTHVHD-RNGIMTTLNRQGSIVQAKAVDPVAVNFKKVDLKTGPVHDPFSTRM 399

OY 312 DVE-YDVCPTFLGGRDWPSPALNPDGSIYFIPILNNVCYDMAVDOEFTSMOYNTSNVTK 370

Db 400 DHKGNICPSAMGPHNGGLADYDPDRTYFGLNHCIDMEWEPMLPYRACGFVFGATILAM 459

OY 371 LP-----FGKDWIGRIDIDISTRTILMSVERAANYSPVLTSGGVLFGNGTDRYFRALS 426

Db 460 YPGNPGFTKEMQVILAMGVGVEYKMTKEKFSYWGSLTKGGLVFYTLIDGNIRKALD 519

OY 427 QETGETLMQTRLATVASGQAISEYDGMQYVAIAGGVSYSYG 469

Db 520 KTNKKEIKFKKMPGSAIGAPMSYAYKGYIA-----TNGVG 557

RESULT 13

OYAGW3 PRELIMINARY; PRT; 623 AA.

AC OYAGW3: 01-JUN-2001 (TRENBLREL, 17, Created)

DT 01-JUN-2001 (TRENBLREL, 17, Last sequence update)

DE 01-DEC-2001 (TRENBLREL, 19, Last annotation update)

OS Pseudomonas butanovora.

OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;

OX NCBI_TaxID=86174;

RP SEQUENCE FROM N.A.

RA Sayavedra-Soto L.A., Vangnai A.S., Arp D.J.:

RT "Characterization of the expression of two distinct alcohol

RT butanovora."

RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF326086; AAK1506.1;

DR HSSP; Q924J7; 1FLG.

DR InterPro; IPR002372; Bac_PQQ repeat.

DR Pfam; PF01011; Bacterial_PQQ; 4.

KW Signal.

FT SIGNAL.

SO SEQUENCE 623 AA; 67553 MW; A573A1C64AC12D55 CRC64;

Query Match 14.7%; Score 371; DB 2; Length 623;
Best Local Similarity 23.3%; Pred. No. 4e-19;
Matches 147; Conservative 75; Mismatches 203; Indels 206; Gaps 24;

OY 2 KPTSLMSAGALALLAFAFAQVPTVDLTLAN--PRAGEISIGQOENYRHSPLTOI 59

Db 7 KFLAIRAIVATAALSLPA-AAVDTYMEDIANDHKTGTGDLVTLGLKAKRHSPLKAI 65

OY 60 TTEVNGQLOLWVA-----RGMQPKQVQVPLIHGVMY----- 92

Db 66 NTDVNAVLPVAMSFSGEKGQOE---AQV--LVHDDVIYATASYSIRFAIDARSKRL 120

OY 93 -----LANIVANGVIVAGSTCOYSPFCFSGHSDATGELMRNTF----- 133

Db 121 WEINARLPDDIRPCDVNRGAATYGVFETLDAAMVALDRKTGKVVRKKEFGDHKYG 180

OY 134 -----IPRAGEEDDETW----- 145

Db 181 YTMGAPVYIKOKSGRILLVHGSSGDFGVVGLFARDPDTGEVWARPNWEGHGRIN 240

OY 146 GND-----YEARMMWTA---WQIITDPTNLVHGSTAVGPA 180

Db 241 GKSTFTGPKAPSWPDDPNSPTGKVEA--WSGGGAPWOTASFDVNNNVYIGAGNPAPW 299

OY 181 SETORTG-----GTLXGNTFPAVRPTGEIIVRHQTLPRDMDQECFEMVNTN-- 232

Db 300 NTKRTAPGDDPRRWDSLF-TSGQAVDASTGELKGFYQIHPNDAMPDSGNNSVYLEIK 358

OY 233 -----VDVOPSTEMEGLOSIN-----PNATGGERLVLTGCKTGMQ--F 272

Db 359 DPKTKKVVNSAHADHNGFFEVTDRLMLAKAGIPIKFTS---LIGAMPVDDITVASGF 415

OY 273 DAETGEFLMAKDTNQNMIESTIDENGITVNEADILTELDEYDVCPFTLGGDRWPSAL 332

Db 416 DLTKGRI-EKDRPPQPEKAGDKGESIFV-----PFLGGNNHPIKSY 459

OY 333 NPDGSIYFIPILNNVCYDMAVDOEFTSMOYNTSNV-----KLPPG-----KDMIG 379

Db 460 SPDTGLEIIPAHMMAMD-----YMEENTYKAGSAVLDQGRINLFDHVG 506

OY 380 RIDAIDISTRTILMSVERAANYSPVLTSGGVLFGNGTDRYFRALSQETGETLMQTRLA 439

Db 507 ILRAIDPSPARSLAGQGVPA-VAGTLITTAGGVFTGSDGYLAKAPDAKNGKELMKRQTG 565

OY 440 TVASGQAISEYDGMQYVAIAGGVSYSYG 470

Db 566 SGVSVVPTWEMDEQYVAIISG---YGVAV 593

RESULT 14

OY34G0 PRELIMINARY; PRT; 695 AA.

AC OY34G0: 01-DEC-2001 (TRENBLREL, 19, Created)

DT 01-DEC-2001 (TRENBLREL, 19, Last sequence update)

DE 01-DEC-2001 (TRENBLREL, 19, Last annotation update)

OS Pseudomonas sp. DH2001.

OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;

OX NCBI_TaxID=163360;

RP SEQUENCE FROM N.A.

RA Hopper D.J., Kaderbhai M.A., Little A.R., Mariotti S.A., Young M.,

RA Rogozinski J.:

RT "Cloning, sequencing and analysis of the gene for lupanine

RT hydroxylase, a quinoxalochrome c from a Pseudomonas sp."

RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AJ318095; CAC67410.1;

KW Signal.

FT SIGNAL.

FT CHAIN

SO SEQUENCE 695 AA; 74956 MW; 9096C6387E457FE0 CRC64;

Query Match 14.4%; Score 364.5; DB 2; Length 695;
Best Local Similarity 24.2%; Pred. No. 1.4e-18;
Matches 130; Conservative 67; Mismatches 198; Indels 143; Gaps 19;

OY 38 AGEWISYIGQOENYRHSPLTOITTEVNGQLOLVVARMOGKQV--TPRIHGWYTL--- 93

Db 36 SGWMSLLGGGNGDHYSALAKDVNKSNNVKNLGSFTMEKGDGLVNPGLADGVITGCGP 95

OY 94 -ANTVANGV----- 109

Db 96 PKGIYANDLKTGNLMTVTPVOYDKDTSMTGWTGHVNRGLAVDDNVYISGYLLAV 155

OY 110 -----YSPFGC-----FVS-----GH-----DSATGEL 128

Db 156 SPTTKLITWSSQCDPKKMOAITGAPRVGGKVFIGNASGDRGGRGLDPAFKTKGRL 215

OY 129 WRNFIPTPAGE-----EGDEFWGNDEYARMM-----TGAWGQITVDPVTNLVHY 172

Db 216 WREYTMP--GDPSKPFENDLLAKASTKGTIDY---WKYTKGVSPWAIITVDEASDTLYF 270

OY 173 GSTAVGPASETQGTGGLYGTNTRFAVRPTGEIIVRHQTLPRDMDQECFEMVNTN 232

Db 271 GTDPSFWSPAQRAPDAPGDELFSHIIAVDASTGAYKWHFOTVONDGNSNSATVHTMLAD 330

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd

OM protein - protein search, using sw model

Run on: May 24, 2002, 10:19:16 ; Search time 101.42 Seconds
(without alignments)
549.784 Million cell updates/sec

Title:	WALICK-934-125..PEP
Perfect score:	2689
Sequence:	1 MKPTSLMASGALALIAAP.....GMOYVAIAGGVSTGSGLNS 502

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched:	747574 seqs, 111073796 residues	747574
Total number of hits satisfying chosen parameters:		

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing:	Minimum Match 0%
	Maximum Match 100%
	Listing first 45 summaries

```
Database :
1: A.Genesepc_0328072.*
2: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1980.DAT.*
3: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1981.DAT.*
4: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1982.DAT.*
5: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1983.DAT.*
6: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1984.DAT.*
7: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1985.DAT.*
8: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1986.DAT.*
9: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1987.DAT.*
10: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1988.DAT.*
11: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1989.DAT.*
12: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1990.DAT.*
13: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1991.DAT.*
14: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1992.DAT.*
15: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1993.DAT.*
16: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1994.DAT.*
17: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1995.DAT.*
18: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1996.DAT.*
19: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1997.DAT.*
20: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1998.DAT.*
21: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1999.DAT.*
22: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA2000.DAT.*
23: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA2001.DAT.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2652	98.6	579	19	AAM37873	Alcohol and/or ald
2	2327	86.5	579	19	AAM37874	Alcohol and/or ald
3	2304.5	85.7	579	19	AAM37876	Alcohol and/or ald
4	2233.5	83.1	578	19	AAM37875	Alcohol and/or ald
5	476	17.7	754	21	AAB35887	Sorbitol dehydroge
6	474.5	17.6	738	13	AAR20192	Sorbitol dehydroge
7	474	17.6	742	11	AAR05335	ADH complex protei
8	473.5	17.6	738	12	AAR13893	Amino acid sequenc
9	307.5	11.4	740	20	AAM935019	A.alloceitigenes me
10	153	5.7	443	20	AAB95019	Sorbitol dehydroge
11	140.5	5.2	715	22	ABC24430	Novel human diagno
					AAG81738	S. epidermidis ope

12	140	5.2	910	22	AAG83007	S. epidermidis ope
13	139.5	5.2	676	22	AAG82914	S. epidermidis ope
14	136.5	5.1	470	22	AAG82701	Novel human diagno
15	129	4.8	958	22	AAG82509	Novel human diagno
16	129	4.8	1510	22	ABG21573	Protein G variant
17	126.5	4.7	593	11	AA607014	Staphylococcus aur
18	126.5	4.7	1016	22	AA6034349	Sequence of polype
19	125.5	4.7	480	8	AA670948	Pathogenic Staphyl
20	125.5	4.7	824	18	AA6709614	Protein G variant-
21	119	4.4	269	11	AA607006	IgG-binding Strept
22	119	4.4	269	11	AA607006	Streptococcus prot
23	118.5	4.4	593	15	AA662945	Streptococcus prot
24	118	4.4	269	10	AA694788	Protein G variant
25	116	4.3	594	12	AA692772	S. epidermidis ope
26	114.5	4.3	404	22	AA685225	Streptococcus Gx78
27	112.5	4.2	514	22	AA685285	S. epidermidis ope
28	110.5	4.1	278	22	AA682635	Human protein sequ
29	110.5	4.1	501	18	AA610645	S. epidermidis ope
30	109.5	4.1	228	21	AA610433	HPV6 mutant LI pro
31	108.5	4.1	501	18	AA610644	Expression vector
32	108.5	4.0	501	18	AA610644	HPV6 mutant LI pro
33	108.5	4.0	501	18	AA626109	HPV11 mutant LI pr
34	108.5	4.0	530	22	AAU12236	HPV11 mutant LI pr
35	107.5	4.0	501	18	AA626107	Human PRO340 poly
36	107.5	4.0	770	22	AA630824	HPV11 mutant LI pr
37	107.5	4.0	998	10	AA637706	Amino acid sequenc
38	107.5	4.0	998	10	AA637706	Sequence of the an
39	106	3.9	649	22	AA633621	Eleimera tenella 45
40	106	3.9	2659	22	AA686530	Drosophila melanog
41	105.5	3.9	429	22	AA6865426	Drosophila melanog
42	105.5	3.9	429	22	AA6693799	Human interferon-a
43	105.5	3.9	429	22	AA6693799	Human protein sequ
44	105.5	3.9	4659	21	AA656578	Human prostate can
45	105.5	3.9	501	21	AA69981	HPV 6 LI protein m
			996	18	AA633624	Eleimera tenella 45

ALIGNMENTS

XX	Asakura A,	Hoshino T,	Ojima S,	Shin'ich M,	Tomiyama N;
XX	(HOFF)	HOFFMANN LA ROCHE & CO AG F.			
PA	19-SEP-1996;	96EP-0115001.			
PR	11-SEP-1997;	97EP-0115801.			
XX	01-APR-1998.				
PD	EP832974-AZ.				
PN	Protein	"note="mature protein"			
FT	Peptide	1..23			
FM	Key	location/Qualifiers			
XX	Gluconobacter oxydans.				
OS	Alcohol/aldehyde dehydrogenase A enzyme; recombinant Organism				
KW	aldehyde; ketone; carboxylic acid; L-sorbose; D-sorbitol;				
KX	2-keto-L-gulonate acid; L-ascorbic; inhibition.				
DE	Alcohol and/or aldehyde dehydrogenase A amino acid sequence.				
XX	10-AUG-1998 (first entry)				
DT	AAM37873;				
AC	AAM37873 standard; Protein; 579 AA.				
XX	AAM37873				
XX	RESULT 1				
ID	AAM37873				

DR WPI: 1998-195228/18.
 DR N-PSDB: AAV29051.
 XX
 PT Recombinant gluconobacter oxydans alcohol and/or aldehyde
 PT dehydrogenase enzyme(s) - useful for converting substrate(s), e.g.
 XX L-sorbose or D-sorbitol to 2-keto-L-gulonic acid
 PS Claim 1: Pages 35-37; 59pp; English.
 XX
 CC This is the amino acid sequence for the gluconobacter oxydans
 CC alcohol and/or aldehyde dehydrogenase A' enzyme. The enzymes
 CC or recombinant organisms can be used to convert suitable substrates
 CC to aldehydes, ketones or carboxylic acids, especially to convert
 CC L-sorbose or D-sorbitol to 2-keto-L-gulonic acid, which can be
 CC converted to L-ascorbic acid by standard procedures. The derivatives
 CC of ADH enzymes have desired substrate specificity, higher affinity
 CC to a substrate, lower affinity to an inhibitory compound, higher
 CC stability against temperature and/or pH and higher catalytic speed.
 XX
 SQ Sequence 579 AA:

Query Match 98.6%; Score 2652; DB 19; Length 579;
 Best Local Similarity 90.3%; Pred. No. 7.8e-227;
 Matches 502; Conservative 0; Mismatches 0; Indels 54; Gaps 1;
 QY 1 MKPTSLIMASAGALALAPAPAOVTPYTDLLANPAGEMISYGOQENYRHSPLTQIT 60
 DB 1 mkptslilwasagalalapaafaytvtdeilannpagewisygqenyrhsplttqit 60
 QY 61 TENVGOLQIWMARGMOPKQVOTPLIHGQVYLANPGDVYQAIADAKTGDIIMWRHQLPN 120
 DB 61 tenvgqlqlwvargmppkqvotplihdgymylanpgdvaiqaidaktgdlimehrrqlpn 120
 QY 121 IATLN-----I 126
 DB 121 iatlnsfgeptlrmalygttnvfyvswdnhlvaldtatgvtftvdrgggedmwnssgpi 180
 QY 127 VANGYIVAGSTCOYSPFCEYSGHDSATGELMRNFIYPAGEEGDETMNGNDEYRMMTG 186
 DB 161 vanygiavagstcyspficetvsgshdsatgelmrnyfiprageegdetmngndeyrmmtg 240
 QY 187 AMGQITDPVTNLVHGSTAVGPASEFQRTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 246
 DB 241 awgqitdyprtnlvhgstavgpasefqrtpgctgctgctgctgctgctgctgctgctgclp 300
 QY 247 RDNDQCTREEMAVTNDVOPSTEMEGLOSINPNATGERRVLTGVPCKTGEMMOQPAET 306
 DB 301 rdndqctremvntndvopstemeglostinpnatgererrvltgvpcktgctgmqidaet 360
 QY 307 GEFLMARDTNYQNMIESIDENGIVTVNBDALIKELDYEDYDCTPFLGGRDMPAALNPD 366
 DB 361 geflwardtngymiesidengivtneadalkeldyeydvcptflggrdmpaalnpps 420
 QY 367 GIYFELNVCYDMAVDOETSMQVNTSNVTKLPKQKIDIGRIDIDISTGRTTMSV 426
 DB 421 giyfelnvcydmavdoetsmqvntsnvtklppkqkimgrididistgtrtmsv 480
 QY 427 RAAANTSPVLSTGGVLFNGGTRFRALSGEGETIMQRTLATVASGAISYEVDMQY 486
 DB 481 raaantspvletggvlfnggtrfralsgetelmqrtrlatvasgaaisyevdmqy 540
 QY 487 VALIAGGVSYSGIUNS 502
 DB 541 valaggvsvsgisuns 556

RESULT 2
 AAM37874
 ID AAM37874 standard; Protein; 579 AA.
 XX
 AC AAM37874;
 XX

DT 10-AUG-1998 (first entry)
 XX
 DE Alcohol and/or aldehyde dehydrogenase A' amino acid sequence.
 XX
 KW Alcohol/aldehyde dehydrogenase A' enzyme; recombinant organism;
 KW aldehyde; ketone; carboxylic acid; L-sorbose; D-sorbitol;
 KW 2-keto-L-gulonic acid; L-ascorbic; inhibition.
 XX
 OS Gluconobacter oxydans.
 XX
 EH Key Location/Qualifiers
 FT Peptide 1..23
 FT Protein /note="signal peptide"
 FT Protein /note="mature protein"
 XX
 PN EP832974-A2.
 XX
 PD 01-APR-1998.
 XX
 PF 11-SEP-1997; 97EP-0115801.
 XX
 PR 19-SEP-1996; 96EP-0115001.
 XX
 PA (HOFF) HOFFMANN LA ROCHE & CO AG F.
 XX
 PI Asakura A, Hoshino T, Ojima S, Shinjoh M, Tomiyama N;
 XX
 DR WPI: 1998-195228/18.
 DR N-PSDB: AAV29052.
 XX
 PT Recombinant gluconobacter oxydans alcohol and/or aldehyde
 PT dehydrogenase enzyme(s) - useful for converting substrate(s), e.g.
 XX L-sorbose or D-sorbitol to 2-keto-L-gulonic acid
 XX
 PS Claim 1: Pages 38-40; 59pp; English.
 XX
 SQ Sequence 579 AA:

Query Match 86.5%; Score 2327; DB 19; Length 579;
 Best Local Similarity 78.3%; Pred. No. 6.1e-198;
 Matches 432; Conservative 32; Mismatches 34; Indels 54; Gaps 1;
 QY 1 MKPTSLIMASAGALALAPAPAOVTPYTDLLANPAGEMISYGOQENYRHSPLTQIT 60
 DB 1 mkptslilwasagalalapaafaytvtdeilannpagewisygqenyrhsplttqit 60
 QY 61 TENVGOLQIWMARGMOPKQVOTPLIHGQVYLANPGDVYQAIADAKTGDIIMWRHQLPN 120
 DB 61 tenvgqlqlwvargmppkqvotplihdgymylanpgdvaiqaidaktgdlimehrrqlpn 120
 QY 121 IATLN-----I 126
 DB 121 iatlnsfgeptlrmalygttnvfyvswdnhlvaldtatgvtftvdrgggedmwnssgpi 180
 QY 127 VANGYIVAGSTCOYSPFCEYSGHDSATGELMRNFIYPAGEEGDETMNGNDEYRMMTG 186
 DB 181 vanygiavagstcyspficetvsgshdsatgelmrnyfiprageegdetmngndeyrmmtg 240
 QY 187 AMGQITDPVTNLVHGSTAVGPASEFQRTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 246
 DB 241 awgqitdyprtnlvhgstavgpasefqrtpgctgctgctgctgctgctgctgctgctgclp 300

CC	stability against temperature and/or pH and higher catalytic speed.
XX	Sequence 579 AA;
SO	Query Match 85.7%; Score 2304.5; DB 19; Length 579; Best Local Similarity 76.3%; Pred. No. 6; Le-196; Matches 425; Conservative 37; Mismatches 40; Indels 55; Gaps
OY	1 MKPSLMAASGALALLAAPFAQVTPVDELLANPAGEMISTQONENYRHSPLFOIT 60
DB	1 mptcllitsaavlllcapaafgvpitdellaaipagaewinygrnqnyhsplqlgt 60
OY	61 TENVCLOLVMARGMQPKVOVTPPLINDGVMYLANPEDVIOAIDAKTGLIWEHRQLPN 120
DB	61 adhvagllqlvarmaeqaavvtpmlhdyamyianpgdvlgaldagtlgdlwehrrqlpa 120
OY	121 IATLN----- 125
DB	121 valinagdgdkrgvalyftsllyfsawdhllalamefgqvfvdivergsgedjlsnttgp 180
OY	126 IVANGVIVAGSTQVSPBQCIVSGHDSATGELMKNFTIPRAGEGDETWGNDYEARMMT 185
DB	181 ivangvivaagstcgyapyciflsyndsaaqeelwnhltiprgeegdelatvgnidfearmtc 240
OY	186 GAKQITVDPVNTLVHGSFAVGPASSETRGTRPGTILXGTMFRVVRPDNGEIVMRHOTL 245
DB	241 gwvqgltvdpvntlvfygstgvgpasetqrgtprgltlygnlttrlavrpdngelivrngtl 300
OY	246 PRDNDQDCEPRFEMVNTVNDVQPSSTEMEGLOSINPNATGERRVLTGVPCKTGMQDAE 305
DB	301 prdnvdgctfemvnanvdvqpsamegllraimnaatgerrvltgapocktglmstidaa 360
OY	306 TGEPLIARDINTYONNIESIDENGIVTVNEDAILKELDYEDVDCPFLGGRDPSPALNPD 365
DB	361 tgefliaardintymiaasidetgltvnedavlkeldvdcvptllggrdcsaalnpd 420
OY	366 SGITFPLNNVCYDKMAVDOEFTSMDEVNTSNVTKLPPKDMTIGRIDALIDISFGRTMSV 425
DB	421 tgytflpfnacdydlnavdqefsaaldyntsataklapgfemngrridaidsicgrlwsa 480
OY	426 ERAAANSPLVSTGGGVLPNGSTDRFPAALSOEGETLMQRLATVVASGOAISVEYDGMQ 485
DB	481 erpaanspvlstaggvfnngldryflralsgegetlvgarlatavatqaisyeldgvq 540
OY	486 YVAIAGGVSYSGSLNS 502
DB	541 ylaiaaggltygqlna 557
RESULT 4	
AAW37875	AAW37875 standard; Protein: 578 AA.
ID	AAW37875 standard; Protein: 578 AA.
XX	AAW37875;
AC	10-AUG-1998 (first entry)
XX	Alcohol and/or aldehyde dehydrogenase A'' amino acid sequence.
DE	Alcohol and/or aldehyde dehydrogenase A'' enzyme; recombinant organism;
XX	Alcohol/aldehyde dehydrogenase A'' enzyme; recombinant organism;
KW	aldehyde; ketone; carboxylic acid; L-sorbose; D-sorbitol;
KM	2-keto-L-gulononic acid; L-ascorbic; inhibition.
XX	Glucocobacter oxydans.
OS	Location/Qualifiers
XX	Key 1..23
PH	Peptide /note- "signal peptide"
FT	Protein 24..578
FT	/note- "mature protein"
XX	
PN	EP832974-A2.

XX 01-APR-1998.
 XX 11-SEP-1997; 97EP-0115801.
 XX 19-SEP-1996; 96EP-0115001.
 XX (HOFF) HOFFMANN LA ROCHE & CO AG F.
 XX Asakura A, Hoshino T, Ojima S, Shinjoh M, Tomiyama N,
 XX WPI, 1998-195228/18.
 XX N-PSDB; AAV29053.
 XX
 XX Recombinant Gluconobacter oxydans alcohol and/or aldehyde
 XX dehydrogenase enzyme(s) - useful for converting substrate(s), e.g.
 XX L-sorbose or D-sorbitol to 2-keto-L-gulonic acid
 XX
 XX Claim 1; Pages 41-43; 59pp; English.
 XX
 XX This is the amino acid sequence for the Gluconobacter oxydans
 XX alcohol and/or aldehyde dehydrogenase A' enzyme. The enzymes
 XX or recombinant organisms can be used to convert suitable substrates
 XX to aldehydes, ketones or carboxylic acids, especially to convert
 XX L-sorbose or D-sorbitol to 2-keto-L-gulonic acid, which can be
 XX converted to L-ascorbic acid by standard procedures. The derivatives
 XX of ADH enzymes have desired substrate specificity, higher affinity
 XX to a substrate, lower affinity to an inhibitory compound, higher
 XX stability against temperature and/or pH and higher catalytic speed.
 XX
 XX Sequence 578 AA;

Query Match 83.18; Score 2233.5; DB 19; Length 578;
 Best Local Similarity 75.18; Pred. No. 1.2e-189;
 Matches 417; Conservative 37; Mismatches 46; Indels 55; Gaps 2;

QY 1 MKPTSLMASHAGALALALAFPAFAOVPTDELAMPAGEWISTGONENRHSPLQITL 60
 DB 1 mktllqssaaallvltipalag-taltdemlanppggwinyqgenyrsplqtlt 59
 QY 61 TENVGQLVWARGMQPGKVQVTPILHDCVWLANPGDVQALDAKGDILMHRQLPN 120
 DB 60 adnvgqqlvwarqmeagklytvlphdgvmylanpgdvaldaatgdlwetrqlpn 119
 QY 121 IATLN-----I 126
 DB 120 iatlnsfgeptgmalylgtnyfvswdhvaldtsqgvvfdvrggtdmvsnsqpl 179
 QY 127 VANGVIVAGSTCOYSPFCFVSGHDSATGEELMRNFITPRAGEGDETMNDYERAKMTG 186
 DB 180 vanyvfvagstcqvfpfcfvsghdsatgaelwrnffipragegdetwngdyetrmtg 239
 QY 187 AWGQITDTPVNVLYHYSSTAVGPASETORCTPGCTLGTNTREAVRPTGETVWRHQTLP 246
 DB 240 wvgqitdtpvngvlyhysstavgpaasetgrgtvsgmyntrctfavrptgetvwrhqtlp 299
 QY 247 RDNDDECTEMVNTVDVOPSTMEGLSINPMNATSERVLTGVPCKTGMQFAET 306
 DB 300 rdnddectemvntvdvopstmeeglshainpdaatgervltgvpckngtmqfaet 359
 QY 307 GEFLMADTNYONMIESIDENGIVYVNDALIKELDYEDVCPTEFLGRMPAALNPDS 366
 DB 360 geflmdtndyongmiesidengivvndalikelddyedvcpflegrpmppaalnpds 419
 QY 367 GTFIFLNNVCYDMAVDEFTSMQVNTSVTKLPKGMIRDAIDISTGRTLMASVE 426
 DB 420 gtfiflnnvcydmavdeftsmqvntsvtklpgkmirdaidistgtrtlmasve 479
 QY 427 RAANANSPVLSGGVLENGSTDRFRALSOETGETIMQRLATVASGAISYEDVGQY 486
 DB 480 ryasnospvlsggvlengstdrfralsetgetimqrlatvasgaaisyedvgqy 539

QY 487 VALAGGVSYSGSLN 501
 DB 540 valagrvtsysgsln 554

RESULT 5
 AAB35987
 ID AAB35987 standard; Protein; 754 AA.
 AC AAB35987;
 XX
 XX 01-MAR-2001 (first entry)

DE Sorbitol dehydrogenase subunit 1 amino acid sequence.
 XX
 XX Sorbitol dehydrogenase; SDH; Gluconobacter oxydans; D-sorbitol;
 XX L-sorbose production; 2-keto-L-gulonic acid.
 XX
 XX Gluconobacter oxydans.
 XX
 XX MO200065066-A1.
 XX
 XX 02-NOV-2000.
 XX
 XX 23-APR-1999; 99WO-IB00736.
 XX
 XX 23-APR-1999; 99WO-IB00736.
 XX
 XX 23-APR-1999; 99WO-IB00736.
 XX
 XX (CHOI/) CHOI E.
 XX (RHEE/) RHEE S.
 XX (LEE/) LEE E.
 XX
 XX Choi E, Rhee S, Lee E;
 XX
 XX WPI; 2000-687351/67.
 XX N-PSDB; AAC83153.

PT Novel membrane-bound sorbitol dehydrogenase nucleic acid molecule
 PT isolated from Gluconobacter suboxydans useful for the fermentative
 PT production of 2-keto-L-gulonic acid and L-sorbose from D-sorbitol
 PS
 PS Claim 1; Fig 8; 96pp; English.

CC This invention relates to an isolated membrane-bound sorbitol
 CC dehydrogenase (SDH) from Gluconobacter oxydans. The invention includes
 CC polynucleotides (AAC83153 - AAC83154) encoding the 3 subunits of SDH
 CC (AAB35987 - AAB35989). Also included in the invention are two
 CC coding regions. The sorbitol dehydrogenase polynucleotide sequences are
 CC useful for producing L-sorbose from D-sorbitol and for increasing the
 CC production of 2-keto-L-gulonic acid by transforming a host cell,
 CC especially Gluconobacter with the DNA and selecting the transformed host
 CC genes) encoding fragments of SDH sequences (fragments of the SDH subunit
 CC sequences are not given separately in the specification but are included
 CC in sequences AAC83156 and AAC83157.
 XX
 XX Sequence 754 AA;

Query Match 17.7%; Score 476; DB 21; Length 754;
 Best Local Similarity 26.3%; Pred. No. 3.1e-33;
 Matches 158; Conservative 80; Mismatches 230; Indels 132; Gaps 18;

QY 14 LALILAAPFAOVTPV-----IDELANPAGEMWISTGONENRHSPLQITTEWVGQ 66
 DB 17 lgcacaaiafctspvalaetdgtatlnadqhgdmvsygrtysqiryspldqltkdaasn 76
 QY 67 LQVWAKMQPGKVQV-TPLIHGVAVYLANPGDVQALDAKGDILMHRQLPN-NIA-- 122
 DB 77 lqlawhydltdnrggegtplivdgvmyatnwmkaldatgkllvsvdkpkyvniadt 136
 QY 123 -----TLNIVA--NGVIVAGS-----TCQTS 141


```

137 gcoedvnmgaaywnkvfyfgtfdgrllaldaktkylwsvytwpkkaqlghnrsyvdga 196
142 P-----FGC--FVSGHDSATGEELMRYFIPRAGEGD----- 172
197 prlakgkvligngaefargvrtaydaetgmwvffitvnpnkpddgaasddvmska 256
173 -ETWGDYEFARMWTG---AMGQITPDVYTLVHKGSTAVGPASTQGTGCTLYGNR 228
257 yptwagagawkgqgsgvrtwdsliydpytdlvylygnspwnkfkfseggnllfgsl 316
229 FAVRPDTGEIWRHQTLPDRNDQCEFEEMWYTVQVSTEMEGLOSINFNATGE-RR 287
317 valnptdkgyvwhfgetpmdgwtysvqgimalamp-----ngemrh 359
OY 288 VLVGVPCKTGMQFPAETGEELMARDTYQMIESIDE-NGIYVNDALIKELDEYD 346
Db 360 vlvnap-kgnfilyldaktkgkfkisgkpylenwangldpvtgrnynpdalwtlmgkpw 418
OY 347 VCPFLGGRDMPASALNPDSGITFIFLNNVCY-----DMAVDQEFSTMDVY 393
Db 419 gipdligghmfaamayspqtklvyipaqqvpyvdpqkgfkahdswmlgldmkiql 478
OY 394 NTSN---VTKLPCKMIGRIDAIDISTGRILMSVERAANYSPLSTGGGVLFNGGTD 449
Db 479 ddnqgpkhdaqgfkldkglwivawdpqkgaafvvdhkgpwnngllacagvllfglan 538
OY 450 RYFRALSOETGLMOTRLATVASGOAISYEVDMQVAT---AG-----GVSYSGS 499
Db 539 gefhaydattgkdlitfrpaqallappvtylangkqyvaevwgsllypfflgyartsg 598

RESULT 6
AAR20192
ID AAR20192 standard; Protein: 738 AA.
XX
XX AAR20192;
AC
XX
XX 16-APR-1992 (first entry)
DT
XX
XX ADH complex protein (mol.wt. 72.000).
DE
XX
XX Alcohol dehydrogenase: acetic acid; fermentation.
KM
XX
XX Acetobacter alioaceti: genes NH-24.
OS
XX
XX JP03266988-A.
PN
XX
XX 27-NOV-1991.
PD
XX
XX 26-MAR-1990; 90JP-0073440.
PF
XX
XX 26-FEB-1990; 90JP-0042301.
PR
XX
XX 26-MAR-1990; 90JP-0073440.
PA
XX
XX (NAKA-) NAKANO SUTEN KK.
DR
XX
XX WPI: 1992-019325/03.
DR
XX
XX N-PSDB: AAQ20383.
PT
XX
XX Alcohol dehydrogenase complex structural gene - used in plasmid
PT
XX
XX and enhancing efficiency of acetic acid fermentation for
PT
XX
XX transformed acetic acid bacteria
PS
XX
XX Disclosure: Fig 3(1-3); 21pp; Japanese.
XX
XX Acetobacter transformed with the sequence encoding this protein can
XX
XX enhance the efficiency of acetic acid fermentation. The ADH complex
XX
XX can be easily extracted from the bacteria and purified and it can be
XX
XX used for the determination of an alcohol.
XX
XX See also AAQ20383-84, and -86-88.
XX
XX
XX Sequence 738 AA:
SQ

```

```

Query Match 17.6%; Score 474.5; DB 13; Length 738;
Best Local Similarity 25.6%; Pred. No. 4e-33;
Matches 151; Conservative 77; Mismatches 229; Indels 133; Gaps 15;

OY 10 SAGLALLAPAFQVTPYVDE-----LIAPPAGEMISYGCNENRHSPLQTIT 61
Db 17 taglcaallsyvatmasaddggatgatgealinhaddpynmlytgsdqyspdlqnr 76
OY 62 ENVGQILDVWAKQKQCKVQV-TPLIHDGVMTLANPGDVIQADAKTGLIWEHRQLP 119
Db 77 snvgnlklawyldldtnrggeqplvdygwyattmwmkavdaatgkllwysqrvp 136
OY 120 NLA-----TN----- 125
Db 137 nhadggcoedvnmgaaywnkvfyfgtfdgrllaldaktkylwsvntlppaealqkrsy 196
OY 126 -----IVANGVIYAGSTCOYSPRC--FVSGHDSATGEELMRYFIPRAGEGD----- 172
Db 197 tvdgaprlakgrvllgn--ggsefgargvsaifaetgkvdwvffvtpnkpneadaas 254
OY 173 -----ETWGDYEFARMWTGA-----WGQITPDVYTLVHKGSTAVGPASTQ 215
Db 255 vlmnkaygtws-----ptgawtrgsgvrtwdslydvpadlvlygngspwnkyxr 307
OY 216 GTPGGTLYGTNTRFAVRPDTEGIWRHQTLPDRNDQCEFEEMWYTVQVSTEMEGIQ 275
Db 308 segkgnllfgslvalkpketgcyvwhfgetpmdqwdfssdqimllidpl----- 357
OY 276 SINPNMATGERRVLGVPCKTGMQFPAETGEELMARDTYQMIESIP-ENGIVTNE 334
Db 358 -----ngeltrivharhknngfilyldaktkgkfkisgkpylenwangldpvtgrnynp 410
OY 335 DAILKELDEYDVCPTFLGGRDMPASALNPDSGITFIFLNNVCYDMAVDQEFSTMDVY 393
Db 411 dalyltqkewyigipgdlqghmfaamayspqtklvyipaqqvpyvdpqkgfkahdswmlgldmkiql 470
OY 394 NTSN---NWTKLPPG-----KDMTGRIDAIDISTGRILMSVERAANYSPLSTGGGV 442
Db 471 nlgldmkiqkvgipdspeakgaafvvdhkgpwnngllacagvllfglan 530
OY 443 LFNQGTDRYFRALSOETGLMOTRLATVASGOAISYEVDMQVAT---AG-----GVSYSGS 492
Db 531 lfglangefhaydattgkdlitfrpaqallappvtylangkqyvaevwgsllypfflgyartsg 598

RESULT 7
AAR05235
ID AAR05235 standard; Protein: 742 AA.
XX
XX AAR05235;
AC
XX
XX 04-AUG-1990 (first entry)
DT
XX
XX Amino acid sequence of alcohol dehydrogenase (ADH) encoded by bases
DE
XX
XX 1-2229 of alcohol dehydrogenase (ADH) gene.
DE
XX
XX Alcohol dehydrogenase (ADH) gene: Acetobacter pasteurianus IFO 3191;
KM
XX
XX Acetobacter aceti K1006 (FERM-7528); plasmid pAA721.
KM
XX
XX Acetobacter aceti strain K1006 (FERM-7528).
OS
XX
XX JP02000452-A.
PN
XX
XX 05-JAN-1990.
PD
XX
XX 30-OCT-1987; 87JP-0273190.
PF
XX
XX 30-OCT-1987; 87JP-0273190, JP-075069.
PR
XX
XX (QPPP ) QP CORP (QPPJ-).
XX
XX
XX
XX

```

DR WPI; 1990-047990/07.
 DR N-PSDB: AAQ91811.
 XX
 PT Cloning DNA, plasmid and microbe contg. it -
 PT contains alcohol dehydrogenase gene derived from Acetobacter
 PT aceti K10006 strain
 XX
 PS Disclosure: Fig 4; 8pp; Japanese.
 XX
 CC Also new are a recombinant plasmid contg. its encoding DNA, and a
 CC microorganism transformed with the plasmid.
 CC A DNA fragment was obtd. from A. aceti
 CC strain K1006 and was transferred to an ADH-defective strain. The
 CC resulting plasmid, PA4721, contg. its gene was inserted into A.
 CC pasteurianus strain IFO 3191 by the triparent method. Its gene is
 CC useful for improving Acetobacter culturing.
 XX
 SQ Sequence 742 AA;

Query Match 17.6%; Score 474; DB 11; Length 742;
 Best Local Similarity 26.0%; Pred. No. 4.5e-33;
 Matches 160; Conservative 87; Mismatches 202; Indels 166; Gaps 23;

OY 11 AGALALIAAFAPVPTVDLL--ANPPAGEWISYCONQENYRHSPLTQITTEWGOLO 68
 DB 23 aaalpyaavparaagagntgeaithadhpemwisygrtyseqrpsldqinsvngdlk 82
 OY 69 LVNARMGMOGKVO-VPLIHDGVNVLNPGDVIAIDAKTGDLIWEHRQLP-NIA----- 122
 DB 83 llyytlidnrgaeatpivdgmjyattwskmealaadagklwqydpkpsnldkqg 142
 OY 123 --TLN----- 126
 DB 143 cdlvngagaymgkvfgtfdgrlvaadaktgkvwavntlpadaalqgrsytlvgav 202
 OY 127 VANGVIVAGSTQYSPFC--FVSGHDSANGELMRYFIPRAGEGD----- 172
 DB 203 vakgylvllgn--ggaefargfvaafdaegklwrfytpnknepdhaadnlmka 260
 OY 173 -ETWGNDEYARMWT-----GAMQITVDPTNIVHGSTAVGASATQNGTGCGLY-- 223
 DB 261 yktwpr--kgaavrvqggsgtwdslvydpsdlly--lavv-----ngspwmykyrse 309
 OY 224 --GTV-----TRFAVRPTGELVWRHQTLPEDNWDQCTEEMVTVNDVOPSTEMEGLSI 277
 DB 310 gigenlflgslvalkpetgeyvwahfqaclpmdqwdylsvqgintldmpvk----- 358
 OY 278 NPNAATGE-RRVLTVGVCCKTGMOPDAENGELIAROTNYQNMLESID-----EN 327
 DB 359 -----gemhvlvhar-kngffylidaktgeflsgknygyvgnwngldp1tgrpmynd 411
 OY 328 GIWTVNEDALIKELDEYDVCPEFLGGRDPSALNPDGFIYFIPLNVCV----- 378
 DB 412 gilytlng-----kfwygl--ppglgahfiamayspkthlvypahgiprpykngyggf 463
 OY 379 -----DMMAVDOETSMDEVYNTSNVTKLPKCKMDIGRIDAIDISGRILWSERAAANYSP 434
 DB 464 kphadswngyldmtkngrpdtpre-artayikdlngwllawbrvkmekvwwkldhngpng 522
 OY 435 VISTGGGVFNGCTDRYRRAISQETGETLMOTRLATVAAAGAISEYEVDDGQGYVA----- 488
 DB 523 llatcgddllftgllangefhaydatngsdllykfdaegslilppmtylsvngkqyvaevgw 582
 OY 489 ----IAGGVSYSYSGC 499
 DB 583 gilypsmgvgvrtsg 597

RESULT 8
 AAR13993 standard; Protein: 738 AA.
 XX

AC AAR13993;
 XX
 DT 09-DEC-1991 (first entry)
 XX
 DE A.alcoetigenes membrane-bound ADH 72kd sub-unit.
 XX
 KM alcohol dehydrogenase complex; carboxylic acid production.
 XX
 OS Acetobacter alcoetigenes.
 XX
 PN EP448969-A.
 XX
 PD 02-OCT-1991.
 XX
 PF 26-FEB-1991; 91EP-0102793.
 XX
 PR 26-MAR-1990; 90JP-0073440.
 XX
 PR 26-FEB-1990; 90JP-0042391.
 XX
 PA (NAKA-) NAKANO VINEGAR KK.
 XX
 PI Tamaki T, Takemura H, Tayama K, Fukaya M, Okumura H;
 PI Kawamura Y;
 DR WPI; 1991-289462/40.
 XX
 DR N-PSDB: AAQ13580.
 XX
 PT Gene for membrane-bound alcohol dehydrogenase complex - obtd.
 PT from Acetobacter alcoetigenes, used for prodn. of enzyme for
 PT converting alcohol to acid
 XX
 PS Disclosure: Fig 3; 36pp; English.
 XX
 CC Total DNA was prepared from A.alcoetigenes MH-24, digested with
 CC PstI and SmaI and ligated to PstI-SmaI cleaved pUC18. The ligation
 CC mixture was used to transform E.coli JM109. Probes were designed
 CC based on the N-terminal amino acid sequence of the ADH complex
 CC isolated from A. alcoetigenes (see AAQ13582-013584). The gene
 CC encoding the 72kd ADH subunit was isolated and sequenced. The
 CC directly sequenced N-terminal region of the purified 72kd sub-unit
 CC corresponds to the sequence beginning at residue 36 of the deduced
 CC sequence. This suggests that the first 35 N-terminal amino acids
 CC form a leader peptide involved in secretion of the mature 72kd
 CC protein. The deduced amino acid sequence has 77 per cent homology
 CC with the same enzyme from A. aceti K6033. See also AAQ13581.
 XX
 SQ Sequence 738 AA;

Query Match 17.6%; Score 473.5; DB 12; Length 738;
 Best Local Similarity 25.6%; Pred. No. 4.9e-33;
 Matches 151; Conservative 76; Mismatches 230; Indels 133; Gaps 15;

OY 10 SAGLALIALAPAFAPVPTVDL-----LANPPAGEWISYCONQENYRHSPLTQITT 61
 DB 17 taglcaallsgyatlmassaddggatgatgeaithadhpnmwlygrtyseqrpsldqin 76
 OY 62 ENWGQLOLVNARMGMOGKVOV-TPLIHDGVNVLNPGDVIAIDAKTGDLIWEHRQLP- 119
 DB 77 snvgnlklawlyldlnrgsgqgtplivdgvmvattwsmmkavadaagkllwsydpvpg 136
 OY 120 NIA-----TLN----- 125
 DB 137 nldakgcddtvtvrgaaymgkvfgtfdgrllaldaktgklvsvntlpaealqgrsy 196
 OY 126 ----IVANGVIVAGSTQYSPFC--FVSGHDSANGELMRYFIPRAGEGD----- 172
 DB 197 tvdgapriakgrvllgn--ggsefargfvaafdaegkvdwrfitypnknepdaas 254
 OY 173 -----ETWGNDEYARMWTGA-----MGQITVDPTNIVHGSTAVGASATQNGTGCGLY 215
 DB 255 vlmnkaygtws-----plgawtrvgggtvwdslvydpvalvlylgvngspwmykyr 307

Query Match	11.4%	Score 307.5	DB 20	Length 740	
Best Local Similarity	21.6%	Prod. No. 2.8e-185	Indels 269	Gaps 333	
Matches 148	Conservative 74	Mismatches 195			
DB	23	AAQVTPYTDLLANPRAGCEWISGQNGEYRSHSPLOTQTTENNWSQQL-VMAWG--QPG 78			
DB	86	sqypamapqgsanparqgdwaygridnqitryspiseltpenaskllvalvayhtyngsyrrp 145			
DB	79	KV---QVTP-L-IHDGVMTLANPGDVIQALDITGLIMENH-----RQLNINMT--L 124			
DB	146	qynkwaetctplkvvggllytcsammdilk-lpdratqgkrlrnvdvkynsiptaaqgv 204			
DB	125	NIVANGVIAGSTCOVSPGCFVSGH-----DSATGGEIMKNYFIPRAG----- 169			
DB	205	tyftssvvpqgqch---nrlliegtldmrllavdaetgd-----fcfnfignqavalm 254			
DB	170	-----EGDPTW-----GNDYEARMWTGAM----- 188			
DB	255	gqjgmsvypgfusmtappvngvvvvhveidqgrwapsvngivgydaesgkfvawdwn 314			
DB		-----FAVRD 234			

```

Query Match                               11.4%: Score 307.5: DB 20: Length 740:
Best Local Similarity 21.6%: Pred. No. 2.8e-18:
Matches 148: Conservative 74: Mismatches 195: Indels 269: Gaps 33

OY 23 AOVTPVTDLLANPFRAGEMISGYGQENYRHSRSPLOTTETENVGQQL--VMARGH--QPG 78
DB :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
86 sqvpamapqgsanpargdwagvygiddhqtspiseltpeaasklkvaylhtysgrip 145
OY 79 KV-----QVTP-L-IHDGVMTLANPDDVIGQALDAKDTGSLIMEHR-----KQLNINAT--L 124
DB :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
146 qynkwaetctplkvvggilytcsamndilk-lapattgqivrrnvdkynsiptyackgv 204
OY 125 NIVANGVIAGSFCOXSFGCFVSGH-----DSATGELMRNFIIRACG----- 169
DB :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
205 tyftssvvpogqpcch--nrlllegllamrlaydaetgld-----fcqpfngqgvnl 254
OY 170 -----EGDETW-----GNDYERAMWTGMW-- 188
DB 255 gglgesvpgfytsmtappvlnqvvvvnhveildqgrwapsgvirlydaesgkftwawdn 314
OY 189 -----GQITVDVTNLVHGYSTAVGPASETQKCTPGSLXGNTNR-----FAVRPD 234
DB :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
315 nsgrsqpey-rvtvltaverripqlpdrtrg-----gstllpdrnsaadyssalred 366
OY 235 -----TGEIWRHQTLPDRMDQECTFEEMATVNDVQSTEMEGLSINP 279
DB 367 aenkysavvaldvktgsprvrtfakdkwdyldigsqatl-----mdmpg-----p 413
OY 280 NNAATGERRVLTVGPCKTGTMMQFDATEGTEFLMARDTNYQMIESIDEN-----GIV----- 330
DB :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
414 dsgtlypallm--pkrgqftvladrtrgk-----pllveetpapspvlpgdp 459
OY 331 -----TYNEDATL-----LKELD-----VEY 345
DB 460 rsptqpwsvampallrvpdlkctedmngmspidqlfctrikfranyvgelfrpsvdkpwley 519
OY 346 DWCEPTFGGRDMPSAALNPDSGIVRIFPLN--NVCDDMAVQOEFTSM----- 390
DB 520 ---pnygsgdwsmsydpqsglllanwllcpmydqlytckkadsjglimpddpnfkpg 576
OY 391 -----DYNTSNVTKLPPGKDMIGRIDAIDISNG--RTLWS-- 424
DB 577 ggaegnangmdtptyglvlpfwdqy-tgmncnrip---ygmftalcmkngqkvlwqhp1 631
OY 425 -----VERAANYSPVLTSGGVLENG--GDRYFRALSOETGELMOTR 467
DB 632 gtarangpwwlpqglpwlwlyctpnngsvvftgsgllfisaatdqiraidentkvwvasv 691
OY 468 LATVASGCAISIEVDGMOYVAIAGGG 493
DB :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
692 lpgggqanpmtyeanghnyavalmag 717

RESULT 10
ABG24430
ID ABG24430 standard; Protein: 443 AA.
AC ABG24430;
DE 18-FEB-2002 (first entry)
DB Novel human diagnostic protein #24421.
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX food supplement; medical imaging; diagnostic; genetic disorder;
XX Homo sapiens.
XX W0200175067-A2.
XX 11-OCT-2001.
XX

```

XX	Dimanac RT, Liu C, Tang YT
DR	WPI; 2001-639362/73.
DR	N-PSDB; AAS88617.
XX	

new isolated polynucleotide and encoded polypeptides, useful in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits and to assess biodiversity -

Claim 20; SEQ ID No 54789; 103pp; English.

the invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical disorders involving aberrant (II). (I) and (II) are useful for treating disorders involving aberrant polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations, and to produce other types of data and products dependent on DNA and amino acid sequences. Abc00010-Abc30377 represent novel human diagnostic amino acid sequences of the invention.

Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at [ftp.wipo.int/pub/published_pct_sequences](http://wipo.int/pub/published_pct_sequences).

Query Match	5.7%;	Score 153;	DB 22;	Length 443;
Best Local Similarity	30.8%;	Pred. No. 7e-05;		
Matches 44:	Conservative			

	Accession	Size	Indels	Gaps
QY	6 L1M4S-----AGALALLAARPAQVPTVIDELIANPAGIEWISYQONENRHSPLNO	58		
Db	168 ltwagfrdppeingtclisadatpaeas-lspvadq-----dwpaygrnqeggrfspiqlq	218		
QY	59 ITTENTNGQLQLYWA-----RGMQPKGV--QVTPRLIHDCVMNLIANGDYTAIDKTKTD	109		
Db	219 inadynhnlkeawfrtgdvkqpnodpsaithevrlpvgtllylctahrrlialdaasqk	278		
QY	110 L1WEHRROL-PIMLATLIANYANGV	131		
Db	279 ekhydpelktneisfghwtvcrgv	301		

RESULT	11
AAAG81738	
ID	AAAG81738 standard; Protein; 715 AA
XX	

XX
DT 03-SEP-2001 (first entry)
XX

epidermids open reading frame protein sequence SEQ ID NO:570.

<p><i>Staphylococcus epidermidis</i> SRI strain; infection; diagnosis; vaccination; endocarditis.</p>	<p>100</p>
---	------------

XX	Staphylococcus epidermidis.
OS	
XX	W0200134809-A2.
PN	
XX	17-MAY-2001.
XX	
XX	09-NOV-2000; 2000MO-US30782
PF	
XX	09-NOV-1999; 9905-0164258
PR	
XX	(GLAX) GLAXO GROUP LTD.
PA	
XX	Kimmerly W0;
PI	
XX	
XX	
DR	WPI: 2001-316495/33.
	N-FSDB; AAH52588.

Nucleic acids encoding polypeptides from *Staphylococcus epidermidis* useful for vaccinating against infections, e.g. endocarditis -
Claim 18; Page 189; 2188pp; English.

CC (II'), given in AAmS3970 to AAmS3970 represent nucleic acids (II) encoding polypeptides
CC (II) and (II') can have antibacterial activity and therefore can be used
CC in vaccination. The nucleic acids (I) may be used to produce the
CC S. epidermidis polypeptides (II) via the production of vectors
CC containing them which are used to produce hosts cells which express the
CC polypeptides. The polypeptides (II) (and/or nucleic acids) may then be
CC used to vaccinate subjects and to raise antibodies against the bacteria
CC activity and therefore identify compounds that may be used for the
CC treatment of S. epidermidis infections, e.g., endocarditis. AAmS3971 to
CC AAmS3990 represent specifically claimed S. epidermidis genomic DNA
CC polynucleotide sequences from the present invention. AAmS3991 to
CC AAmS5098 represent oligonucleotide sequences and primers which are used
CC in the exemplification of the present invention.
CC N.B. The present invention specifically claims all the polynucleotide
CC sequences given in the sequence listing of the present specification,
CC however the sequence listing only goes up to SEQ ID NO:4454 so even
CC though sequences are given in the disclosure for SEQ ID NO:4455 to 4472,
CC no sequences are present for SEQ ID NO:4455 to 4464.
CC
CC Sequence 715 AA:
CC
CC

Query Match	5.2%;	Score 140.5;	DB 22;	Length 715;
Best Local Similarity	23.8%;	Pred. No. 0.0019;		
Matches 105				

```

20 1PVIDELLANPAGEWISYSGONQENYRHSPLTOIT-----TENWGOLQVARGMQ- 76
309 tptt---knpitqekvagepfevkfnmvdvibvuzgocibh---bbaaa-:||

```

365 dypgkpgvknpdtgevtpprvdvtkygvpddpi-----tsieipfdkkrffglan 430

421 tekvvqg---epvckltlp-----tknplgkv-----gegept----- 455

```
b      |  
456   -----kvtkqpvdeivhygeeikpglhkdefdpnapkgsgqtctgpkpg-----    499  
  
230 AVRDDEGTGRTVWDQDTDDPDPDPAEAE
```

Db 546 kpgervqkgeptkltitptknpitgkvggeptekiktpvde--iteygsgeik 603
 OY 339 KELDEYDYVCPFLAGRWPS--AALNPDSC-IYFIPLANVCY-----DMAVNOE--F 387
 Db 604 pghkdefdpnarksgedvpgkpvknptdgvvtpvddvltkypvddqplsteeelpf 663
 OY 388 TSMQVNTSNVTKLPGRKDM 408
 Db 664 dkrefnpd---lkpgkerv 680

RESULT 12

AAG83007 standard; Protein: 910 AA.

AAG83007; (first entry)

S. epidermidis open reading frame protein sequence SEQ ID NO:3108.

Staphylococcus epidermidis SRI strain; infection; diagnosis;

vaccination; endocarditis.

Staphylococcus epidermidis.

WO200134809-A2.

17-MAY-2001.

09-NOV-2000; 2000MO-US30782.

09-NOV-1999; 99US-0164258.

(GLAX) GLAXO GROUP LTD.

Kimmerly WJ;

WPI: 2001-316495/33.

N-PSDB: AAH53857.

Nucleic acids encoding polypeptides from Staphylococcus epidermidis, useful for vaccinating against infections, e.g. endocarditis.

Claim 18; Page 819; 2188pp; English.

PS AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides
 CC (II), given in AAG81454 to AAG83120, from Staphylococcus epidermidis.
 CC (I) and (II) can have antibacterial activity and therefore can be used
 CC in vaccination. The nucleic acids (I) may be used to produce the
 CC S. epidermidis polypeptides (II) via the production of vectors
 CC containing them which are used to produce hosts cells which express the
 CC polypeptides. The polypeptides (II) (and/or nucleic acids) may then be
 CC used to vaccinate subjects and to raise antibodies against the bacteria.
 CC The polypeptides may also be used to assay for other inhibitors of their
 CC activity and therefore identify compounds that may be used for the
 CC treatment of S. epidermidis infections, e.g. endocarditis. AAH53971 to
 CC AAH55090 represent specifically claimed S. epidermidis genomic DNA
 CC polynucleotide sequences from the present invention. AAH55091 to
 CC AAH55098 represent oligonucleotide sequences and primers which are used
 CC in the exemplification of the present invention.
 CC N.B. The present invention specifically claims all the polynucleotide
 CC sequences given in the sequence listing of the present specification,
 CC however the sequence listing only goes up to SEQ ID NO:4454 so even
 CC though sequences are given in the disclosure for SEQ ID NO:4465 to 4472,
 CC no sequences are present for SEQ ID NO:4455 to 4464.

Sequence 910 AA;

Query Match 5.28; Score 140; DB 22; Length 910;
 Best local similarity 25.48; Pred. No. 0.003; Indels 94; Gaps 23;
 Matches 98; Conservative 45; Mismatches 149;

OY 26 TPVTDLLANPAPGEMISTGONQENYRHSPLFOIT--ENV-----GQLQLVWANGMO- 76
 Db 231 tptt-----knpltgkvggeptekvktqvdeltkygsgeikpghkdefdnarkysge 286
 OY 77 --PGK-----VOVTEPLIHGVMYLANPQDVYQIADAKTGLIMEHNRQL-PNIA-- 122
 Db 287 dvpqkpgvknptdgvvtpvddvltkypvddqplsteeelpfokkrefdnlapg 342
 OY 123 TLNVANGVIVAGSTCOYSPFGCFVSGHDSATGEELMNNYFIPRAGEGDETWGNDYEAR 182
 Db 343 tekvvqg--epqtkltitp-----ltknplgkcv-----gepept----- 377
 OY 183 WMTGAMGOIITYDPTVNLVHGSTAVGPA--SETQSTPGGTILXGNTPRFAVR-PDTGEIV 239
 Db 378 -----ekvtkqpydelvhygsgeikpghkdefdnarkysgedvpgkpvknptdgvv 431
 OY 240 WRHOTLPDNDMDQECTEMMYTN--VDVOSTEMEGL-----QSTNPANAG-ERYVLT 290
 Db 432 -----tpvdd-----vtkypvddqplsteeelpfokkrefdnlapgtekvvqk 477
 OY 291 GVP-GKT-GTMMQFPAETGEFLMARDTNYQMLTSDENGIYVNDALILKELDEYDYVC 348
 Db 478 gepqtkltitptknpitgkvggeptekiktpvde--ivhygsgeikpghkdefdn 535
 OY 349 PTFUGRDWPS--AALNPDSCIYFIP 372
 Db 536 apksqgdvpgkpvknptdgvvtp 561

RESULT 13

AAG82914 standard; Protein: 696 AA.

AAG82914;

03-SEP-2001 (first entry)

S. epidermidis open reading frame protein sequence SEQ ID NO:2922.

Staphylococcus epidermidis SRI strain; infection; diagnosis;

vaccination; endocarditis.

Staphylococcus epidermidis.

WO200134809-A2.

17-MAY-2001.

09-NOV-2000; 2000MO-US30782.

09-NOV-1999; 99US-0164258.

(GLAX) GLAXO GROUP LTD.

Kimmerly WJ;

WPI: 2001-316495/33.

N-PSDB: AAH53764.

Nucleic acids encoding polypeptides from Staphylococcus epidermidis, useful for vaccinating against infections, e.g. endocarditis.

Claim 18; Page 763-764; 2188pp; English.

PS AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides
 CC (II), given in AAG81454 to AAG83120, from Staphylococcus epidermidis.
 CC (I) and (II) can have antibacterial activity and therefore can be used
 CC in vaccination. The nucleic acids (I) may be used to produce the
 CC S. epidermidis polypeptides (II) via the production of vectors
 CC containing them which are used to produce hosts cells which express the
 CC polypeptides. The polypeptides (II) (and/or nucleic acids) may then be
 CC used to vaccinate subjects and to raise antibodies against the bacteria.

yy
os Homo sapiens.

XX
PN W0200175067-A2.

11-OCT-2001.

XX
PF 30-MAR-2001; 2001WO-US08631.

XX
21-MAR-2000: 2000US-0540217.

PR 23-AUG-2000; 2000US-064910 /

PA (HYSE-) HYSEQ INC.

Drmanac RT, Liu C, Tang YT, PI

AA
DR WPI; 2001-639362/73.

DR N-PSDB; AND000000
XX

PT New isolated polymers, gene mapping, identification of mutations
PT diagnostics, forensics, to assess other traits and to assess

PT responsible for genetic disorders - biodiversity

XX
XX
30. CEO ID No 56268: 103pp; English.

to isolated polynucleotide (I) and

polypeptide (II) sequences. (I) is useful as hybridization probes, primers, oligomers, and for chromosome

polymerase chain reaction (PCR) and gene mapping, and in recombinant production of (11). The expressed sequence tags

polynucleotides expressed genes. (I) is useful in gene therapy involving for identifying expressed genes. (I) is useful in gene therapy involving

to restore normal activity or (iii) to generate antibodies against it, detecting or (iv) is useful for generating antibodies against it, detecting or as weight markers and as

quantitating a polypeptide in tissue, as well as its binding partners are useful in medical research and treatment (11) and its binding partners are useful for treating

imaging of sites expressing (II). (I) and (II) are

CC imaging of sites expressing protein expression or biological activity.

The polypeptide and polynucleotide sequences have ap-

CC responsible for genetic disorders or other trials to assess
CC diagnosis, treatment and products dependent on DNA and
CC of data and products dependent on DNA and

CC and to produce cancer cells. ABG00010-ABG30377 represent novel human amino acid sequences. ABG30377 represents a novel human amino acid sequence.

CC diagnostic amino acid sequences in the printed format directly from WIPO.

specification, but was obtained in 100% yield.

XX 6

AA	Sequence	948 AA;
SQ		

Score 129: DB 22: Length 948;

Query Match	29.98;	Pred. No. 0.03;	
Best Local Similarity	17	Matches	30;
		Indels	54;
		Gaps	8;

Matches	43;	conservative	58

[illegible]

Db 310 ltwaqfndpqeintlsadalpaea-ispvaay

59 ITTENGOLOLWMA-----RGMQPGKV--GVIFLIHDOVNIENMZCZ-
QY :|||: |||:
|||

Db 361 inadnvn1keawfrtgdvkgpndpgeltnevtpr-----

00 110 ---LWHRR-----QLPNIAL 124

401 tnljctaharlfa1daemdgiatl 424

Search completed: May 24, 2002, 10:19:18

Fri May 24 11:27:24 2002

walick-934-125.pep.rag

Page 12

Fri May 24 11:27:25 2002

wallick-934-125.pcp.ral

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2002, 10:20:03 ; Search time 39.53 seconds
(without alignments)
310.186 Million cell updates/sec

Title: WALICK-934-125.PCP
2689
Sequence: 1 MKPTSLMMSAGALALALP.....GMOYVATAGSGVSGSLNS 502

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 231628 seqs, 24425594 residues
Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep.*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep.*
3: /cgn2_6/ptodata/2/1aa/5A.COMB.pep.*
4: /cgn2_6/ptodata/2/1aa/5B.COMB.pep.*
5: /cgn2_6/ptodata/2/1aa/5A.COMB.pep.*
6: /cgn2_6/ptodata/2/1aa/5B.COMB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	476	17.7	754	4	US-09-296-284-4
2	474.5	17.6	738	1	US-07-985-458-3
3	467	17.4	720	4	US-09-296-284-25
4	312.5	11.6	739	1	US-09-136-251-2
5	105	3.9	1012	1	US-08-219-262B-10
6	105	3.9	1012	1	US-09-031-655-10
7	102.5	3.8	1042	3	US-08-700-651-5
8	102	3.8	1042	3	US-08-928-361B-11
9	102	3.8	1042	3	US-08-928-361B-11
10	99	3.7	1042	3	US-08-928-361B-11
11	99	3.7	1042	3	US-08-928-361B-11
12	98.5	3.7	1042	3	US-08-928-361B-11
13	98.5	3.7	1042	3	US-08-928-361B-11
14	96.5	3.6	1043	4	US-08-928-361B-11
15	96.5	3.6	1043	4	US-08-928-361B-11
16	96.5	3.6	1043	4	US-08-928-361B-11
17	96.5	3.6	1043	4	US-08-928-361B-11
18	95.5	3.6	1043	4	US-08-928-361B-11
19	95.5	3.6	1043	4	US-08-928-361B-11
20	95	3.5	1043	4	US-08-928-361B-11
21	95	3.5	1043	4	US-08-928-361B-11
22	95	3.5	1043	4	US-08-928-361B-11
23	95	3.5	1043	4	US-08-928-361B-11
24	95	3.5	1043	4	US-08-928-361B-11
25	94.5	3.5	1043	4	US-08-928-361B-11
26	94.5	3.5	1043	4	US-08-928-361B-11
27	94.5	3.5	1043	4	US-08-928-361B-11

28	94.5	3.5	1612	1	US-08-169-927-2	Sequence 2, Appli
29	93.5	3.5	348	1	US-08-247-902A-2	Sequence 2, Appli
30	93.5	3.5	816	1	US-07-731-157A-4	Sequence 4, Appli
31	93.5	3.5	816	1	US-08-229-44A-2	Sequence 2, Appli
32	93.5	3.5	816	2	US-08-341-780-4	Sequence 4, Appli
33	93	3.5	380	2	US-08-971-782-4	Sequence 4, Appli
34	93	3.5	380	4	US-09-309-026-4	Sequence 2, Appli
35	93	3.5	459	3	US-08-971-782-2	Sequence 2, Appli
36	93	3.5	459	4	US-09-309-026-2	Sequence 16, Appli
37	93	3.5	1381	4	US-09-540-245A-16	Sequence 8, Appli
38	92.5	3.4	824	4	US-09-626-589-3	Sequence 10, Appli
39	91.5	3.4	774	4	US-08-462-484-10	Sequence 10, Appli
40	91	3.4	527	1	US-08-441-147-10	Sequence 10, Appli
41	91	3.4	527	5	PCT-US95-07536-10	Sequence 10, Appli
42	91	3.4	527	5	US-08-793-229-32	Sequence 32, Appli
43	91	3.4	551	2	US-09-285-957-32	Sequence 2, Appli
44	91	3.4	551	3	US-09-626-589-2	Sequence 2, Appli
45	91	3.4	659	4	US-09-626-589-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-09-296-284-4
Sequence 4, Application US/09296284A
Patent No. 6204040
GENERAL INFORMATION:
APPLICANT: Choi, Eun-Sung
APPLICANT: Lee, Eun-Hae
TITLE OF INVENTION: Glucoconabacter suboxydans sorbitol dehydrogenase, Genes
FILE OF INVENTION: and Methods of Use Thereof
FILE REFERENCE: 1533.0870000
CURRENT APPLICATION NUMBER: US/09/296,284A
CURRENT FILING DATE: 1999-04-22
NUMBER OF SEQ ID NOS: 87
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 4
LENGTH: 754
TYPE: PRT
ORGANISM: Glucoconabacter suboxydans
US-09-296-284-4

Query Match	17.7%	Score 476	DB 4	Length 754
Best Local Similarity	26.3%	Pred. No. 2.1e-37		
Matches 158	Conservative	80	Mismatches 230	Indels 132
			Gaps	18
QY	14	LALLAARPAQVYV	---	TDLLANPPAGEWISYGNONENRHSPLQITTEVNGO 66
DB	17	LGCAALALFCAISPVALLAEDTGTATINADHPDMMISYGRYSSEQYSPLOITDKMSN 76		
QY	67	LOIWAROMOGKQV	---	TPLIHGVWYLANPGVIOALDAKGLDIMEHRQLP-NIA-- 122
DB	77	ILALHHYDLDINROEGEPLVDGVMTATNWSKALDAATKILMSYDPKVINADR 136		
QY	123	---TNTVA---NCYVAGS---		---TCQYS 141
DB	137	GCDDVNRGAAYVKNKVEFGTDRILALDAKGLKLVSYTYVKEAOLGHRYSYVDA 196		
QY	142	-----FGC-----FVSGHDSATGEELWKRYFIPRAGEED		----- 172
DB	197	PRIAKKVITIGGAEFGARGFVYDAETKMDMREFYVNPDKNDGASADVLMASKA 256		
QY	173	---ETWMDYKRWMTG---AMQITVDPYTNLVHGSTAVGPASFTQRTGPGCLYGTNR 228		
DB	257	YPTWGGGAMKQGGGGGVWDSLITDPTDLVYLGVGSPWNNKFRSEKGNLFLGSI 316		
QY	229	FAVRPDTGELVWRHQLPNDNDQCTEEMAYTNDVQPSFEMGLOSINPNATGE-RR 287		
DB	317	VALNPDTGKYVWHFQETPMDDMDYTSVQOIALDKPV-----NGEMRH 359		

288 VLVGPGCKTGTMMQFPAETGEFLMARDFYONMIESIDE-NGIYTVEDALTELDEYD 346
 Db 360 VLVHAP-KNGFFIIDAKTGKISKPTTYEMWANGDPTGPNYDMLWTLNGKPMY 418
 QY 347 VCTPFGGRDMPSSALNPDSCGYTFPLNNVCY-----DKMAVDOEFTSMAY 393
 Db 419 GIPGDLGGINFAMAYSPQTKIVYTPAQOQVFFVDPQKGFKAHDSMNLGLDMKIGLL 478
 QY 394 NTSN----VTKLPGKMDIGRIDALIDISTGTMTSVERAANSPVLSTGGVLENGCTD 449
 Db 479 DNDNDPOHAKADKQFLKDKGWIYAMDPOKQAAFTVDHKGPMNGGLATAGVLEFOGLAN 538
 QY 450 RFRALSOETGETLMQFALATVASGAQISTEVDGMOYVAL---AG-----GGVSYSG 499
 Db 539 GFHAYDTATGKDLFTPEPAOSAIITAPVYTYANGKQYVAVEVGWGIYTFPLGVAKTSG 598

RESULT 2

US-07-985-458-3
 ; Sequence 3, Application US/07985458
 ; Patent No. 5344777

GENERAL INFORMATION:

APPLICANT: Tamaki, Toshimi;
 APPLICANT: Takemura, Hiroshi;
 APPLICANT: Takemura, Kenji;
 APPLICANT: Fukaya, Masahiro;
 APPLICANT: Okumura, Hajime and
 APPLICANT: Kawamura, Yoshiya
 TITLE OF INVENTION: Structural Gene of Membrane-Bound
 TITLE OF INVENTION: Alcohol Dehydrogenase Complex, Plasmid
 TITLE OF INVENTION: Containing the Same and Transformed Acetic Acid
 NUMBER OF SEQUENCES: 4
 CORRESPONDENCE ADDRESSES:
 ADDRESSEE: Fishauf, Holtz, Goodman & Woodward, P.C.
 STREET: 600 Third Avenue
 CITY: New York
 STATE: New York
 COUNTRY: USA
 ZIP: 10016-2088

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 inch, 0.72 mb
 COMPUTER: IBM PC compatible (NEC PC-9801 ES)
 OPERATING SYSTEM: MS DOS
 SOFTWARE: ASCII Form
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/07/985,458
 FILING DATE: 19921203
 CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/658,221
 FILING DATE: 20-FEB-1991
 APPLICATION NUMBER: 73440/1990
 FILING DATE: 26-MAR-1990
 ATTORNEY/AGENT INFORMATION:
 NAME: Goodman, Herbert

REGISTRATION NUMBER: 17081

REFERENCE/DOCKET NUMBER: 910134/HG
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212)972-1400
 TELEFAX: (212)370-1622

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:
 LENGTH: 738 amino acids
 TYPE: amino acid
 TOPOLOGY: unknown
 MOLECULE TYPE: protein
 FEATURE:
 NAME/KEY: MATURE PEPTIDE
 LOCATION: 36 to 738
 IDENTIFICATION METHOD: N-terminal sequences of the
 IDENTIFICATION METHOD: purified protein having a molecular weight of about

IDENTIFICATION METHOD: 72,000
 ORIGINAL SOURCE:
 ORGANISM: Acetobacter alioacetigenes
 STRAIN: MH-24
 PUBLICATION INFORMATION:
 AUTHORS: Tamaki, Toshimi;
 AUTHORS: Fukaya, Masahiro;
 AUTHORS: Takemura, Hiroshi;
 AUTHORS: Takemura, Kenji;
 AUTHORS: Okumura, Hajime;
 AUTHORS: Kawamura, Yoshiya;
 AUTHORS: Nishiyama, Makoto;
 AUTHORS: Horiuchi, Sueharu and
 Heppu, Teruhiko
 TITLE: Cloning and Sequencing of the Gene Cluster
 TITLE: Encoding Two Subunits of Membrane-Bound
 TITLE: Alcohol Dehydrogenase from Acetobacter
 JOURNAL: Biochimica et Biophysica Acta.
 VOLUME: 1088
 PAGES: 292-300
 DATE: 1991
 US-07-985-458-3

Query Match 17.6%; Score 474.5; DB 1; Length 738;
 Best Local Similarity 25.6%; Pred. No. 2,8e-37;
 Matches 151; Conservative 77; Mismatches 229; Indels 133; Gaps 15;

QY 10 SAGALALLAPAFAYPTIDE-----LIAPPAGEWISYQONENYHSPLOTIT 61
 Db 17 TAGTICAAISGYATVASADDDGAGTGAITHADHPGNMWTYGRYSDDQYSPLDINR 76
 QY 62 ENYGQQLVWARGMPGKQYV-TPLIHGVMYLANPGVIOAIDAKGDLIMEHRLGP- 119
 Db 77 SNVGNLKLAMYLDLDLNRGQEGTPLVIDGVYATTNMSKAAVATGKLLMSIDRVRP 136
 QY 120 NIA-----TLN-----
 Db 137 NIADKCCDVTNRGAAYNNGKYYETGTFDGLIALDAKTGLVMSVMTIPPEAEIKQRSY 125
 QY 126 ----YVANGYVASTGQYSPGCG--FVSGHDSATGEELMNYFLPRAGEGCD----- 196
 Db 197 TVDGAPRIAKGKRVIIIGN--GGSEFGARGFVSAPDAETGKDMKFFVYPPKNEPPAASDS 254
 QY 173 -----ETWGDYEAHMTGA-----WGQITYPVTVLVHYGSTAVPASSETOR 215
 Db 255 VLMNKAYQTNW-----PTGAWTROGGGGTWDSDIVYDPAVLVYLGVNGSPMNYKYR 307
 QY 216 GTPGGLTYGTVTRAPVAPDGTETVWRHQTLPDNDPQECTFEMANTYNDVOPSTEMEGLQ 275
 Db 308 SEGKGDMLFGLSIALKPEFGEYVWHFOETPPDQMDFTSDQIMTLPLT----- 357
 QY 276 SINPNAATGERRYLTVPCKTGTMMQFPAETGEFLMARDFYONMIESID-ENGIVTVNE 334
 Db 358 -----NGETRHVYIHAARKNGFFIIDAKTGEFISGRNYVYVWNASGLDPKTRPIYNP 410
 QY 335 DAIKELDEYDVCPFLGGRDMPSSALNPDSCGYTFPLNNVCYDMAVADQFT-SNDVY 393
 Db 411 DALYTLTGKEMVIGIPDGLGHNFAMAFSPKGLVYIYIAQOVPLLYTYNOYGGFTPHDSW 470
 QY 394 NTSN----NVTKLPGP-----KDMIGRIDALIDISTGTMTSVERAANSPVLSTGGV 442
 Db 471 NLGLDMNKYGIQDSEPAKQAFYKDKGWIYAMDPOKQAAEMRVYDHKGPWNGGLATAGDGL 530
 QY 443 LFNGGIDRYFRALSOETGETLMQFALATVASGAQISTEVDGMOYVALIAG 492
 Db 531 LFQGLANGFHAAYDATNGSDLFHFAADSGIITAPVYTYLANGKQYVAEVEG 580

RESULT 3

US-09-296-284-25
 ; Sequence 25, Application US/09296284A

Fri May 24 11:27:25 2002

walick-934-125.pwp.ra

Page 3

```

? Patent No. 6204040
? GENERAL INFORMATION:
? APPLICANT: Choi, Eun-Sung
? APPLICANT: Rhee, Sang-Ki
? APPLICANT: Lee, Eun-Hae
? TITLE OF INVENTION: Gluconobacter Suboxydans Sorbitol Dehydrogenase, Genes
? TITLE OF INVENTION: and Methods of Use Thereof
? FILE REFERENCE: 1533.087000
? CURRENT APPLICATION NUMBER: US/09/296,284A
? CURRENT FILING DATE: 1999-04-22
? NUMBER OF SEQ. ID NOS. 87
? SOFTWARE: PatentIn Ver. 2.0
? SEQ ID NO 25
? LENGTH: 720
? TYPE: prt
? ORGANISM: Gluconobacter suboxydans
? US-09-296-284-25

```

Query Match	17.4%;	Score 467;	DB 4;	Length 720;
Best Local Similarity	26.2%;	Prod. No. 1.4e-36;		
Matches 152;	Conservative 81;	Mismatches 220;	Indels 128;	Gaps 18;

[illegible]

RESULT 4
US-09-136-251-2
Sequence 2, Application US/09136251A
Patent No. 6127156
GENERAL INFORMATION:
APPLICANT: HOSHINO, Tatsuo
APPLICANT: MIYAZAKI, Taro
APPLICANT: MIYAZAKI, Satsuo
APPLICANT: MIYAZAKI, Masako
APPLICANT: SHINOH, Masako

```

? APPLICANT: TOMIYAMA, NO. 61271561Dumi
? TITLE OF INVENTION: D-Sorbitol Dehydrogenase Gene
? FILE REFERENCE: D-Sorbidal Dehydrogenase Gene
? CURRENT APPLICATION NUMBER: US/09/136,251A
? CURRENT FILING DATE: 1998-08-19
? EARLIER APPLICATION NUMBER: EP 97114432.4
? EARLIER FILING DATE: 1997-08-21
? NUMBER OF SEQ ID NOS: 9
? SOFTWARE: Patentln Ver. 2.1
? SEQ ID NO 2
? LENGTH: 739
? TYPE: PRT
? ORGANISM: Gluconobacter suboxydans
? FEATURE:
? NAME/KEY: SIGNAL
? LOCATION: (1)..(24)
US-09-136-251-2

```

Query Match	11.6%	Score 312.5;	DB: 3;	Length 739;
Best Local Similarity	21.8%	Pred. No. 1,6e+21;		
Matches 149;	Conservative 74;	Mismatches 199;	Indels 269;	Gaps 33

[illegible]

RESULT 5

```

US-08-219-262B-10
; Sequence 10, Application US/08219262B
; Patent No. 5788970
; GENERAL INFORMATION:
; APPLICANT: VAKHARIA, VIKRAM
; APPLICANT: SNYDER, DAVID B
; TITLE OF INVENTION: CHIMERIC INFECTIOUS BURSAL DISEASE VIRUS
; TITLE OF INVENTION: CDNA CLONES, EXPRESSION PRODUCTS AND VACCINES BASED
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT
; STREET: 1755 S. JEFFERSON DAVID HIGHWAY, FOURTH FLOOR
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/219,262B
; FILING DATE: 29-MAR-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 2747-047-27
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 413-3000
; TELEFAX: (703) 413-2220
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1012 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Infectious bursal disease virus
; STRAIN: OH
; US-08-219-262B-10

Query Match
Best Local Similarity 3.9%; Score 105; DB 1; Length 1012;
Matches 84; Conservative 45; Mismatches 130; Indels 120; Gaps 19;

QY 196 VTNLVHYGSTAV-----GPAS-----ETORGPGGTLYGTNTRF 229
DB 1 MTNLMDHTQOIVPFIIRSLMPTTGPASIPDDTLEKHLNSETSTYNLVGDTGSGLIYVF 60
QY 230 AVRPDGTGEIYWRHOTLPD---NMDECTEFEMAVT-----NVDVQPSFE 270
DB 61 PGFP--GSVYGAHYTLQSGSYQFDQ-----MLTAQNLPVSYNCRVLSRSLTVRSSTL 113
QY 271 MEGLOSINP--NAATGERRVLTGVPCKTGTMMQFDETFEFLMARDNTNQNNTESI----- 324
DB 114 PGCVATLNGTINAVT-----FGGSLSEL---TDSYNGMLSATANIN 152
QY 325 DENGIVTVNEDAILKELDVEYDVCTPFLGGRDMSALINPDGSIYFTPLNNVC----- 377
DB 153 DRIGNVLVGEQTVLSTSTYDLSTVRLGD--PIPAAGIDP-----KLATCDDSDRPR 204
QY 378 -YDMAAVQ--EE-----TSMDYNTSNVTKLPPGKDM--GRIDALDITGTRILW- 423
DB 205 VITVADEYFSSQLIPSGVKTTLTFANIDALISLSVGSELLFSQVTHISIEVDVTLIF 264
QY 424 -----SVRAAANYSPLVSTGCVILFN--GGTDRYFALSOETGETLMOTRLATVAS 473

```

```

DB 265 IGFDTGEVTVKAVATDFLTGTNNLVFENLGGFTSETLPITSMKLEVTVYRGGT--A 322
QY 474 GQAISEVDGQMYAIIAG 492
DB 323 GDPISWTVSGTLAVTYIGG 341

RESULT 6
US-09-031-655-10
; Sequence 10, Application US/09031655
; Patent No. 6017759
; GENERAL INFORMATION:
; APPLICANT: VAKHARIA, VIKRAM
; APPLICANT: SNYDER, DAVID B
; TITLE OF INVENTION: CHIMERIC INFECTIOUS BURSAL DISEASE VIRUS
; TITLE OF INVENTION: CDNA CLONES, EXPRESSION PRODUCTS AND VACCINES BASED
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT
; STREET: 1755 S. JEFFERSON DAVID HIGHWAY, FOURTH FLOOR
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/031,655
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/219,262
; FILING DATE: 29-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 2747-047-27
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 413-3000
; TELEFAX: (703) 413-2220
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1012 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Infectious bursal disease virus
; STRAIN: OH
; US-09-031-655-10

Query Match
Best Local Similarity 3.9%; Score 105; DB 3; Length 1012;
Matches 84; Conservative 45; Mismatches 130; Indels 120; Gaps 19;

QY 136 VTNLVHYGSTAV-----GPAS-----ETORGPGGTLYGTNTRF 229
DB 1 MTNLMDHTQOIVPFIIRSLMPTTGPASIPDDTLEKHLNSETSTYNLVGDTGSGLIYVF 60
QY 230 AVRPDGTGEIYWRHOTLPD---NMDECTEFEMAVT-----NVDVQPSFE 270
DB 61 PGFP--GSVYGAHYTLQSGSYQFDQ-----MLTAQNLPVSYNCRVLSRSLTVRSSTL 113
QY 271 MEGLOSINP--NAATGERRVLTGVPCKTGTMMQFDETFEFLMARDNTNQNNTESI----- 324

```

```

Db 114 PGVYALNGINMNT-----FOGISEL-----TDYSYNOIMATANIN 152
OY 325 DENGIVNEDAIKEIDVEYDVCPTFLGRDWPASALNPPDSGIYFIFILNNVC----- 377
Db 153 DKIGVNLVGBEVYLSLPTSYLSYVLGD-PIPAAGLDP-----KIMATCDSDRPR 204
OY 378 -YDMAVDO-EF-----TSMQDYNTSNYTKLPGKMDI-GRIDAIDISTGRTLM- 423
Db 205 VYTVTADEYQFSSQLIPSGVKITLFTANIDALTSLSVGSELIFFSOVTHSEVDVTVF 264
OY 424 -----SVERAANSFVLSTGGVLFN--GQDRYFALSOEGETLMQRLATVAS 473
Db 265 IGFDEYVYKAVATDFGLTGTNNLVFNLGPTSEITPTISMLLEVYTYRGGT--A 322
OY 474 GOAISEYVDGMOYVAIAG 492
Db 323 GDPISWTSGLATVTVGG 341

```

```

RESULT 7
US-08-700-651-5
; Sequence 5, Application US/08700651B
; Patent No. 6015882
; GENERAL INFORMATION:
; APPLICANT: PETERSEN, CAROLYN
; APPLICANT: LEECH, JAMES
; APPLICANT: NELSON, RICHARD, C.
; APPLICANT: GUN, JIRI
; TITLE OF INVENTION: VACCINES, ANTIBODIES, PROTEINS, GLYCOPROTEINS, DNAs AND RNAs
; TITLE OF INVENTION: FOR PROPHYLAXIS AND TREATMENT OF Cryptosporidium parvum
; FILE REFERENCE: 480.19-4(HV) US/08/700,651B
; CURRENT FILING DATE: 1997-08-14
; EARLIER FILING DATE: 1995-04-03
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 5
; LENGTH: 1721
; TYPE: PRF
; ORGANISM: Cryptosporidium parvum
; US-08-700-651-5

```

```

Query Match 3.8%; Score 102.5; DB 3; Length 1721;
Best Local Similarity 20.5%; Pred. No. 1.6; Indels 231; Gaps 37;
Matches 135; Conservative 78; Mismatches 216;

OY 25 VTPTDELLANPPAGEWISYGNQF-----NYRHSPLTQITTEVGOLOLVWARKQP 77
Db 979 IDPTTG-LPENNPTGLHINPTNNMDSFAGAYKXAAVANGIKITDNV-----YGLPV 1029
OY 78 GKVOYTPLIHGVWYLANPDVIOAIDAKTGLIMEHRQOLNIVATLVNA----- 128
Db 1030 GEITLTP-----KDPGSDI-PFNSTGELVADSTGKPLNNSTAGVSKRPGLPPIE 1079
OY 129 --NGVI-----VAGSTQYSP-FGCFVSGHDSAT-----GEEIMRNIFIP--RAGEE 170
Db 1080 DENGINLFDPSTNLPIDGNNQLVNPEFTNSYSGSTGTTKPKPIPVNGGVYVDEFAKQD 1139
OY 171 GDETMGNDYEAHWMTGANGQITVDPVNLVHYGTA--VGPASFTORGFG----- 219
Db 1140 ADK--GKD--GLIVPPTNSIKKDPVNTQYSNITGINIP--ETGKAVIFGSLPGSLNYP 1192
OY 220 -----GLVGTFRFAVRPDTGELIWRHQLLPRDN--WQECTFEHMYTN-- 262
Db 1193 SFNPTQDEITIKPVDTYTGLPYDPSGTEITIDPAKILPIGSAVAGEILFEVNTTDE 1252
OY 263 ----VDVQ--PSTMEGLQS-----INP-----NAATG--ERRAVLT 290
Db 1253 VTGPIIDLEGLPRDPVSGILPOLPNGTIVDPNSKKPIGSHSGFINGTSGEOSHEDKDPST 1312

```

```

OY 291 GVPCKTGT-XMOFP-----AETGEFLMARDTNY-----ONMIES 323
Db 1313 GKPLDPNTGLHFPDEDSGLINLPETGDKLOGSHSGFEMPVPGKPOGENGIMTPQOILEA 1372
OY 324 IDENGIVTNE-----DAI-----LKEIDVEYVCTFLAG----- 354
Db 1373 L--NKLIPTSENVNISPRSSDAVDPRPTNWMNKISQOTYQVDSKTTILSASVIFHAL 1430
OY 355 ---RDWPSAL--NPDSGIYFIFLNNVCYDMAVDO-----EFTSMQDYNTSNYTKLPP 403
Db 1431 GPTQDPTTGLPSPDSTGLPPIPGFNVLPDQGEOLKGSVPYSLYVEKKNITVEAY 1490
OY 404 GKMDIGRIDAIDISTGRTLMVERAANSFVLSTG-----GGVLFNGGDRYFRA-- 454
Db 1491 GL-----PVDEKGTGFPIDPI-----STLPPAKNGELIDPISKYSFGSIAGISKAG 1538
OY 455 ---LSOETGETLMOT-----RLATVASGOAISEYVDGMOYVAIAG-----GGV 494
Db 1539 SOSKSDSGNPIDPSTNMPYDPKGGKLLDPESGALIDNSYSGV-FATVPTAARKGGV 1597

```

```

RESULT 8
US-08-928-361B-11
; Sequence 11, Application US/08928361B
; Patent No. 6071518
; GENERAL INFORMATION:
; APPLICANT: Petersen, Carolyn
; APPLICANT: PEPTIDES, POLYPEPTIDES, GLYCOPROTEINS,
; TITLE OF INVENTION: THEIR FUNCTIONAL MUTANTS, VARIANTS, ANALOGS AND FRAGMENTS
; TITLE OF INVENTION: FOR TREATMENT AND DETECTION/DIAGNOSIS OF CRYPTOSPORIDIUM
; TITLE OF INVENTION: SPECIES INFECTIONS
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PETERS, VERNY, JONES & BIKSA
; STREET: 385 Sherman Avenue, suite 6
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306-1840
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/928,361B
; FILING DATE: 12-SEP-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/026,062
; FILING DATE: 13-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: VERNY, Hana
; REGISTRATION NUMBER: 30,518
; REFERENCE/DOCKET NUMBER: 480.76-1(HV)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-324-1677
; TELEFAX: 650-324-1678
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1042 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-928-361B-11

Query Match 3.8%; Score 102; DB 3; Length 1042;
Best Local Similarity 20.9%; Pred. No. 0.79; Indels 232; Gaps 36;
Matches 138; Conservative 71; Mismatches 219;

```

```

QY 25 VTPVTELLANPAGEWISYQNOE-----NYRHSPLQITTEENVGOLQVWARGMOP 77
DB 301 IDPTTG-LTFNPPTGHLINPTNNNTMSSRFAGATKAYVNSGKITDNYVGLPVDITGLPK 359
QY 78 GKQVTPPLHDSGYMLANPCVIOAIDAKTGLIMEH-----RQLPNI-----121
DB 360 DPVSDIP-----FNSITGEL-----VDPSTGKPIINNTAGIVSGKGLPIEDENGFLD 409
QY 122 --ATLNIYANGVI-----VAGST--COYSPGCFVSGHDSATGEELMRYEFLPRAG 168
DB 410 PSTKLPIDGNQOLVNPETNSVSGSTSGSKPKP--GIYVNGGCVVDEE-----AK 459
QY 169 EGEDEMGNDYEARMGAMQITTYDVTNLVHYGSTA--VGPASETORGTPG-----219
DB 460 DQADK--GKD--GLIYPPNINSIKDPVNTQYSNTGNIINP--ETGKVIPLGSLPGSLN 512
QY 220 -----GTYGNTFRFAVRPDTGEIWMRHQTLPRDN--WDQCTEEMAVTN 262
DB 513 YPSENTPOQDEITGKRPVDTVGLPYDPSTGELIDPATKLPISGSVAGDELITVLELNTT 572
QY 263 -----VDVQ-----PSTMEGLQS-----INP-----NAATG-----ERRV 288
DB 573 DEVTGLPDLLEGLPRDPVSGLEPLNGTLVDPNSKKPIPSHSGFLNGTSGQSHKDP 632
QY 289 LTGVPECKTGMQFD-----AETGEFLWARDNTYQNM--ESIDENGIYVNEADAIL 338
DB 633 STGKPLDPNTGLPFDEDSGSLINPETGDKLQSGHSGTFMVPKPGQENGIMTPBQ--IL 691
QY 339 KELDV-----EXDVCP-----TFL--GGRDWPSSAL-----362
DB 692 EALNKLPTSNENINISPRSSDAVDRPTNTMWNKISQOTFOVQGGKITPGSAAVITHAL 751
QY 363 -----NPDGCIYFIPLANVCYDMAVDO-----EFTSMDYNTNSNTKLP 403
DB 752 GTPPTODPTTGLPSDPSGTGLPIPGFVNLVDPQGTQIGSVPIYSLYKEKNITVEAY 811
QY 404 GKDMIGRIDAIDISTGRTMVSVERAANYSPLVSTG-----GVLNFGTDTRYFRA---454
DB 812 GL-----PVDPKGTGFPIDPI-----STLPFAKNGLIDLPISGYFGSGLAGFISGAG 859
QY 455 -----LSOETGETLMOT-----RLATVASGOAISYEVGOMQYVALAG-----GV 494
DB 860 SQSSSDSGNPIIDPSTNMPYDPKTKGLIDPESGIALDINSVSGV--FAIVPGTAAPKKGV 918

RESULT 9
US-08-928-361B-5
; Sequence 5, Application US/08928361B
; Patent No. 6071518
; GENERAL INFORMATION:
; APPLICANT: Petersen, Carolyn
; TITLE OF INVENTION: PEPTIDES, POLYPEPTIDES, GLYCOPROTEINS,
; TITLE OF INVENTION: THEIR FUNCTIONAL MUTANTS, VARIANTS, ANALOGS AND FRAGMENTS
; TITLE OF INVENTION: FOR TREATMENT AND DETECTION/DIAGNOSIS OF CRYPTOSPORIDIUM
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PETERS, VERNY, JONES & BIRKA
; STREET: 385 Sherman Avenue, Suite 6
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306-1840
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PC-DOS/MS-DOS
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/928,361B
; FILING DATE: 12-SEP-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:

```

```

APPLICATION NUMBER: US 60/026,062
FILING DATE: 13-SEP-1996
ATTORNEY/AGENT INFORMATION:
NAME: VERNY, HANA
REGISTRATION NUMBER: 30,518
REFERENCE/DOCKET NUMBER: 480,76-1(HV)
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-324-1677
TELEFAX: 650-324-1678
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1837 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-928-361B-5

```

```

Query Match      3.8%; Score 102; DB 3; Length 1837;
Best Local Similarity 20.98%; Pred. No. 2;
Matches 138; Conservative 71; Mismatches 219; Indels 232; Gaps 36;

```

```

QY 25 VTPVTELLANPAGEWISYQNOE-----NYRHSPLQITTEENVGOLQVWARGMOP 77
DB 1096 IDPTTG-LTFNPPTGHLINPTNNNTMSSRFAGATKAYVNSGKITDNYVGLPVDITGLPK 1154
QY 78 GKQVTPPLHDSGYMLANPCVIOAIDAKTGLIMEH-----RQLPNI-----121
DB 1155 DPVSDIP-----FNSITGEL-----VDPSTGKPIINNTAGIVSGKGLPIEDENGFLD 1204
QY 122 --ATLNIYANGVI-----VAGST--COYSPGCFVSGHDSATGEELMRYEFLPRAG 168
DB 1205 PSTKLPIDGNQOLVNPETNSVSGSTSGSKPKP--GIYVNGGCVVDEE-----AK 1254
QY 169 EGEDEMGNDYEARMGAMQITTYDVTNLVHYGSTA--VGPASETORGTPG-----219
DB 1255 DQADK--GKD--GLIYPPNINSIKDPVNTQYSNTGNIINP--ETGKVIPLGSLPGSLN 1307
QY 220 -----GTYGNTFRFAVRPDTGEIWMRHQTLPRDN--WDQCTEEMAVTN 262
DB 1308 YPSENTPOQDEITGKRPVDTVGLPYDPSTGELIDPATKLPISGSVAGDELITVLELNTT 1367
QY 263 -----VDVQ-----PSTMEGLQS-----INP-----NAATG-----ERRV 288
DB 1368 DEVTGLPDLLEGLPRDPVSGLEPLNGTLVDPNSKKPIPSHSGFLNGTSGQSHKDP 1427
QY 289 LTGVPECKTGMQFD-----AETGEFLWARDNTYQNM--ESIDENGIYVNEADAIL 338
DB 1428 STGKPLDPNTGLPFDEDSGSLINPETGDKLQSGHSGTFMVPKPGQENGIMTPBQ--IL 1486
QY 339 KELDV-----EXDVCP-----TFL--GGRDWPSSAL-----362
DB 1487 EALNKLPTSNENINISPRSSDAVDRPTNTMWNKISQOTFOVQGGKITPGSAAVITHAL 1546
QY 363 -----NPDGCIYFIPLANVCYDMAVDO-----EFTSMDYNTNSNTKLP 403
DB 1547 GTPPTODPTTGLPSDPSGTGLPIPGFVNLVDPQGTQIGSVPIYSLYKEKNITVEAY 1606
QY 404 GKDMIGRIDAIDISTGRTMVSVERAANYSPLVSTG-----GVLNFGTDTRYFRA---454
DB 1607 GL-----PVDPKGTGFPIDPI-----STLPFAKNGLIDLPISGYFGSGLAGFISGAG 1654
QY 455 -----LSOETGETLMOT-----RLATVASGOAISYEVGOMQYVALAG-----GV 494
DB 1655 SQSSSDSGNPIIDPSTNMPYDPKTKGLIDPESGIALDINSVSGV--FAIVPGTAAPKKGV 1713

RESULT 10
US-07-642-734C-4
; Sequence 4, Application US/07642734C
; Patent No. 5824513
; GENERAL INFORMATION:

```


Fri May 24 11:27:25 2002

walick-934-125.pcp.ra1

Page 7

APPLICANT: Katz, L
APPLICANT: Donadio, S
APPLICANT: McAlpine, J B
TITLE OF INVENTION: Recombinant DNA Method for Producing
TITLE OF INVENTION: Erythromycin Analogs
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Edward H. Gorman
STREET: Abbott Laboratories D377/AP6D-2 One Abbott
STREET: Park Rd
CITY: Abbott Park
STATE: IL
COUNTRY: US
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/642,734C
FILING DATE: 17-JAN-91
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Danckers, Andreas M
REGISTRATION NUMBER: 32652
REFERENCE/DOCKET NUMBER: 4952.US.01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 708-937-9396
TELEFAX: 708-938-2623
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 3567 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-642-734C-4

Query Match 3.7%; Score 99; DB 2; Length 3567;
Best local similarity 21.9%; Pred. No. 12;
Matches 122; Conservative 49; Mismatches 211; Indels 174; Gaps 25;
QY 21 AFAQVPTVDLHNPAGW-----ISYQNEVNRHSPLTQTTENVGQLQVW 71
DB 2911 AFAPIAVTDHRLARVDPG-WSDADAAVPIAY-----TTAHVALHDLGLRAGOSVLIH 2964
QY 72 ARGMPQKVQVYPLIHGVMYLANPG----- 97
DB 2965 AAGGVGMAAVLARRAGAEVLATAGPRAKGTILRALGLDDEHIASSRETFARKFRERTG 3024
QY 98 ----DVIQAIIDAKTGLIWEHRQLPNTIYVANGVIV-AGST-----COYSPE 143
DB 3025 GRGVYV--LNSLTGLIDES-----ADL-LAEDGVFVEKGTDLRDAGDFRGYAPF 3074
QY 144 GCFVSGHSAETGELMRYNFIPIRAGEGDETWGNDYFARNMTGAW-----GOITYDPYTNL 199
DB 3075 DLGKAG--DRLGELIRVYVGLGAGELD-----RLPVSAWELGSAPRALQHSRSG 3123
QY 200 VHYGSTAV-GRASETORGT---PGCTLVGTNTRFAVRPDTGEIWRHQTJL-PRDNMDQEC 254
DB 3124 RHVGKLVLTGPAFAPDPDPTGLITGGT--GTGLRLARHLVTEHGVRRHLLVSRGADAPG 3181
QY 255 TFEAMVTVVDVQSTEMEGLOSINPNAAT---GERRVLTGSPCKTGTMOQFDAGTEGFL 310
DB 3182 SEELRAETIEDGASAEIACCTADRDALSLALDGLRPLTGVAAGVYL----- 3230
QY 311 WARDNTYQNTIESIDENGTIT-VNEDALTELKELDEYDVCPTELGGRWPSAALNDGSGIT 369
DB 3231 ----ADLVTSIDEPAYEQLAKYDAA-----VWILHELTANTGLS 3267
QY 370 FIFLNNVCTDMAVVDQEFSTMDYVNTSNV-----TKLPPGK----- 405

DB 3268 FEVLFSASAVIA-----GPGGVYAAANESLNAALALRRREGCLPAKLGMLMAQASEMT 3323
QY 406 ----DMIGRIDAIDISTGTLMSVERAANSPVLSTGGVLFNGGTDYRFRALSOETGE 461
DB 3324 SGLGDRIAR-----TGVAALPETERALALFDSALRGEVVEFLINSALRRAEVPE 3376
QY 462 TL-WOTRLATVYASGOA 476
DB 3377 VLKGMVRAKLRAAGOA 3392

RESULT 11
US-08-439-009A-4
Sequence 4, Application US/08439009A
Patent No. 6004787
GENERAL INFORMATION:
APPLICANT: Donadio, S
APPLICANT: Katz, L
APPLICANT: McAlpine, J B
TITLE OF INVENTION: Method of Directing Biosynthesis of
TITLE OF INVENTION: Specific Polypeptides
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Steven F. Weinstein
STREET: Abbott Laboratories D377/AP6D-2 One Abbott
STREET: Park Rd
CITY: Abbott Park
STATE: IL
COUNTRY: US
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/439,009A
FILING DATE: 11-MAY-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Casuto, Dianne
REGISTRATION NUMBER: 40,943
REFERENCE/DOCKET NUMBER: 4952.US.D1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 847-938-3137
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 3567 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-439-009A-4

Query Match 3.7%; Score 99; DB 3; Length 3567;
Best local similarity 21.9%; Pred. No. 12;
Matches 122; Conservative 49; Mismatches 211; Indels 174; Gaps 25;
QY 21 AFAQVPTVDLHNPAGW-----ISYQNEVNRHSPLTQTTENVGQLQVW 71
DB 2911 AFAPIAVTDHRLARVDPG-WSDADAAVPIAY-----TTAHVALHDLGLRAGOSVLIH 2964
QY 72 ARGMPQKVQVYPLIHGVMYLANPG----- 97
DB 2965 AAGGVGMAAVLARRAGAEVLATAGPRAKGTILRALGLDDEHIASSRETFARKFRERTG 3024
QY 98 ----DVIQAIIDAKTGLIWEHRQLPNTIYVANGVIV-AGST-----COYSPE 143
DB 3025 GRGVYV--LNSLTGLIDES-----ADL-LAEDGVFVEKGTDLRDAGDFRGYAPF 3074
QY 144 GCFVSGHSAETGELMRYNFIPIRAGEGDETWGNDYFARNMTGAW-----GOITYDPYTNL 199
DB 3075 DLGKAG--DRLGELIRVYVGLGAGELD-----RLPVSAWELGSAPRALQHSRSG 3123

```

QY 200 VHGSTAV-GPASETORG-----PGTLYGNTNFAVPDGTGLVWRHQT-L-PRDMWDOEC 254
DB 3124 RHVGKLVLPAPVDPDGTGLTGT--GTLGRLAHLVTEHGVAHLLVSRRADPG 3181
QY 255 TFEWMTVNDVOPSTEMEGIOSINPNAT-----GERRVLTGVPCKTGTWOPDAETGFL 310
DB 3182 SPELAEIEDGASAEIACPTADRLSALLDGLPRPLHGVAAVGL----- 3230
QY 311 WARDNTYONMIESIDENGIVT-VNEDAIILKELDEYDVCPTFLGGRDPSAALNPDGTY 369
DB 3231 -----ADGLVTSIDEPRAVEQVLRKVDA-----WNLHELANTGLS 3267
QY 370 FIPINNYCYDMAVDOETFSMDVYNTSNV-----TKLPPGK----- 405
DB 3268 FFLVFSASAVLA-----GREGGYAANESLUNALALARTGIDPAKALGMLAQAASEMT 3323
QY 406 -----DMIGRIDALIDISTGTLMSVERAANYSPLYSTGGVLENGGTDYFRALSOETGE 461
DB 3324 SGLGDIRAR-----TGVAALPTRRALALDPSALRRGGEVFEPLINSALRRAEFVPE 3376
QY 462 TL-WOTRLATVASGA 476
DB 3377 VLNGWVRAKLRAGA 3392

```

```

RESULT 12
US-08-987-519-2
Sequence 2, Application US/08987519
Patent No. 5952216
GENERAL INFORMATION:
APPLICANT: Luderer, Steven
TITLE OF INVENTION: Synthetic HPV16 Virus-Like Particles
FILE REFERENCE: 19853
CURRENT APPLICATION NUMBER: US/08/987,519
CURRENT FILING DATE: 1997-12-09
NUMBER OF SEQ ID NOS: 3
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 2
LENGTH: 500
TYPE: PRF
ORGANISM: Human Papillomavirus
US-08-987-519-2

```

```

Query Match
Best Local Similarity 3.7%; Score 98.5; DB 2; Length 500;
Matches 56; Conservative 30; Mismatches 88; Indels 89; Gaps 12;
QY 36 PRAGEMISYGONOENYHSPLTQITTEENVGQLOLVWARGMOPGKYQVPLIHDEVMILAN 95
DB 158 PRLGHEWKGKQCTN--TP-----VOAGDCPPLLELITSVIODEDM----- 196
QY 96 PGDIVQALDAKTGDLIWEHRROLPNATLNIYANGV--IVAGSTCYQ-----SPFG- 144
DB 197 -----VDTGFAM-----NEADLOTKSDVPIDICGTCKPDVLOMAADPYGD 240
QY 145 -----CEVSGHDSATGELMNYEIPRAGEGDETWGNDYEAWMTGAMGQITTYDPVNLVH 201
DB 241 RLFFFL-----RKQMFARHFFENRAGEVGPV-----PDLIHK 274
QY 202 YGSTAAGPASEFROGTGTLTGNTNFAVPDGTGLVWRHQTLPDN--WDQECTEM 258
DB 275 GSGNRTSVASSTLYVTPTGSLVSEADLPKPF-----YLOKQAGHNNICWGN-----OL 325
QY 259 WATNVNDVOPSTEMEGIOSINPNA 281
DB 326 FVTVDTRSTNMTLCASTVTTSS 348

```

```

RESULT 13
US-08-868-786-2
Sequence 2, Application US/08868786

```

```

Patent No. 5998701
GENERAL INFORMATION:
APPLICANT: Kawchuk, Lawrence M.
APPLICANT: Armstrong, John
APPLICANT: Lynch, Dermot
APPLICANT: Knowles, Richard
TITLE OF INVENTION: Potatoes Having Improved Quality
NUMBER OF SPOUNCES: 10
CORRESPONDENCE ADDRESS:
ADDRESS: Greenlee, Winner and Sullivan, P.C.
STREET: 5730 Manhattan Circle, Suite 201
CITY: Boulder
STATE: Colorado
COUNTRY: USA
ZIP: 80803
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
FILING DATE: 04-JUN-1997
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/036,946
FILING DATE: 10-FEB-1997
ATTORNEY/AGENT INFORMATION:
NAME: Sullivan, Sally A
REGISTRATION NUMBER: 32,064
REFERENCE/DOCKET NUMBER: 8-97
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 499-8080
TELEFAX: (303) 499-8089
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 966 amino acids
TYPE: amino acids
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-868-786-2

```

```

Query Match
Best Local Similarity 3.7%; Score 98.5; DB 2; Length 966;
Matches 84; Conservative 56; Mismatches 157; Indels 153; Gaps 18;
QY 25 VTPVTDLANPRAGEMISYGONOENYHSPLTQITTEENVGQLOLV-----WARGM 75
DB 613 VTPPRWIRFCNPPLSAIITKWTCTEDM-----VLKTEKLAELQKPADNEDLONREAK 666
QY 76 QPGKVOVTPILHDGVWYLANGDYQALDAKTGDLIWEHRROLPNATLNIYANGVIYAG 135
DB 667 RSNKIKIVSEFLKEKGIYVNP-DAMFDIOYK--IHEYKROLNT-----FG 710
QY 136 STCYQSPGCEVSGHDSATGELMNYEIPRAGEGDETWGNDYE--RMTGAMGQI 191
DB 711 IYIRKKM-----KEMTAERKTN-FYPRVICFGKAFATYVQAKRIVFIDVAGTI 762
QY 192 TYD-----PYTNLVHGSTAVGPASETOR--GTG-----G 220
DB 763 NHDPEIGDLKYVVPVYNSVAELLIPASDLESHISTAGMEASGTSNMKFMAMGCIQIG 822
QY 221 TLVGTNFAVRDPDGTGLIWRHQTLPDNDQECTFEMVYTNVNDVOPSTEMEGIOSINPN 280
DB 823 TLDGANVE--IREVEG-----
QY 281 AATGERRVLTGVPCKTGTWOPDAETGFLAARDTNTONMIESIDENGIVTNEADILKE 340
DB 838 -----ENFFLEGAHBIAGLRKERADGKFV--PDREFEVKEFVRSGARFSYNYDGLIS 891
QY 341 LD-----VEDVCPPTLGLGRDPSAALNPDGTYFIPLNVCYDMAVDOETFSMDVY 393

```

Fri May 24 11:27:25 2002

walick-934-125.pep.rai

Page 9

DB 892 LEGNGFGADY-----FLVGKDFPS-----YIEOEKVDYERKOKRTTMSIL 936
QY 394 NTSNTKLPKPKDM-----IGRIDAIDIS 417
DB 937 NTAGSYKSSDTHIEYAKDINWIEVAEIA 966

RESULT 14
US-08-931-608A-5
Sequence 5, Application US/08931608A
Patent No. 6302655

GENERAL INFORMATION:
APPLICANT: Iobell, Peter
APPLICANT: Sleat, David E.
TITLE OF INVENTION: NOVEL HUMAN LYSOSOMAL PROTEIN AND METHODS OF ITS USE
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: David A. Jackson, Esq.
STREET: 411 Hackensack Ave, Continental Plaza, 4th Floor
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/931,608A
FILING DATE:

CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq, David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 601-1-077
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-343-1684
TELEFAX: 201-487-5800

INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 635 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHEICAL: NO
US-08-931-608A-5

Query Match 3.6%; Score 96.5; DB 4; Length 635;
Best Local Similarity 21.2%; Pred. No. 1.2; Mismatches 245; Indels 131; Gaps 27;
Matches 119; Conservative 65;

DB 1 MKPTSLIMASAGALIALAPAFQVTPVTELLANPPAGEWISYONCINMHSPLTQIT 60
QY 3 IERKALIVAIALAMSSLSAHEADAWSTHTQAMSPASTQV-----LAASSTATT 54
DB 61 TENVGQLOLVARGMQKQVY-----PLHDVWYLANPGDIVQAI-----DAKTG 108
QY 55 TGNVYTLMTGSPRIDAAVTALEADHPLHVEVALKLRN-DALQYFLAGVTPPSALFG 113
DB 109 DLIMEHR---ROLPNIALTLNIVANGIVAGST-CQYSPFCEPVSCHDSANGELMNYFL 164
QY 114 KFLTPSQTERFPGTQGVDAVVAHLQAGFTNIEVAPNKLILSADGTAGAATNGFRITSI 173
DB 165 PRAGEGDEWNGDYEAAMWTGANGQ-----ITYDPVNIWY-----GSTAVGPASE 212
QY 174 KFFSANGREFFANDAPA-LVPSAGDSYNAVGLQNVSKHTLHHYVHPEDVTVPGPNVG 232
DB 213 TQRGTP-----GGTLYGTNTRFAVRPDGTGEIYKHOQTLPRDNWDOCEITEM 259

DB 233 TQAAANAHHHPDEPAIYGGSSILPAATNTAV-----GIITWGSIT----- 273
QY 260 VTNDVDPSTMEGLQSIINPMAATGERVLTGVPCKTGMQEDAE-----TGEFLM 311
DB 274 QTVLDLNSFTSGAGLATVNTTITKVGSGTFANDPDSNGE-WSLDSQDIYVIGAGVKOLIR 332
QY 312 ARDTNVOAMIESIDENGIVT-----VNEDAILKEIDVEYDVCPTPLGGRDMPBALN----- 363
DB 333 YTSANGDSSSGITPDAGITASYNRAVDNIKILNVS-L-----GEDETAQOQSGTQA 384
QY 364 PDGTYFIPLNNVCYDMAVDEFT-----SDVYN-TSNVTKLPKG-KDMIGRIDAIDI 416
DB 385 ADALP-----QQAQAQGTFSIASGAGVQWSTDPISGSRGYANAGTV-KIDL 435
QY 417 STGRTLSYERAAANSP-VLSTGGVLFNGGTDYFRALSOETGETILKQTRATVASQ 475
DB 436 ---THYSSEPAS--SPVIOVGQGTLSST-----TWSGETVWNGLSALPASQ 481
QY 476 AISYVDGMQYVAIAGGVS 495
DB 482 G-----DNNQRLWATGGVS 496

RESULT 15
US-08-928-361B-30
Sequence 30, Application US/08928361B
Patent No. 6071518

GENERAL INFORMATION:
APPLICANT: Petersen, Carolyn
TITLE OF INVENTION: PEPTIDES, POLYPEPTIDES, GLYCOPROTEINS,
TITLE OF INVENTION: THEIR FUNCTIONAL MUTANTS, VARIANTS, ANALOGS AND FRAGMENTS
TITLE OF INVENTION: FOR TREATMENT AND DETECTION/DIAGNOSIS OF CRYPTOSPORIDIUM
NUMBER OF SEQUENCES: 30
CORRESPONDENCE ADDRESS:
ADDRESSEE: PETERS, CARYN, JONES & BIKSA
STREET: 385 Sherman Avenue, Suite 6
CITY: Palo Alto

STATE: CA
COUNTRY: USA
ZIP: 94306-1840

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/928,361B
FILING DATE: 12-SEP-1997

CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/026,062
FILING DATE: 13-SEP-1996
ATTORNEY/AGENT INFORMATION:
NAME: Verry, Hana

REGISTRATION NUMBER: 30,518
REFERENCE/DOCKET NUMBER: 480.76-1(HV)
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-324-1677
TELEFAX: 650-324-1678
INFORMATION FOR SEQ ID NO: 30:
SEQUENCE CHARACTERISTICS:
LENGTH: 1043 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-928-361B-30

Query Match 3.6%; Score 96.5; DB 3; Length 1043;
Best Local Similarity 20.6%; Pred. No. 2.7; Mismatches 231; Indels 231; Gaps 38;
Matches 136; Conservative 76;

```

QY 25 VTPVTDELLANPAGEMISYQONOE-----NYRHSPLTQITTEBNVGOLOLVNARCMOP 77
Db 301 IDPTTG-LPFENPTGHLINPTNNMTDSSFAQAKKAYAVSNGIKTDNV-----YGLPV 351
QY 78 GKVOYTPLIHGVMTLANPGDVIOAIDAKTGLIWEHRROLPTATLNIVA----- 128
Db 352 GETTGLP-----KDPGSDI-PFNSTGTGLVDPSTGKPINNSTAGIVSGKGLPIE 401
QY 129 --NGVI-----VAGSTCOYSP-FCGFSVGHDSAT-----GELMRNYFIP--RAGEE 170
Db 402 DENGTLFDPSTNLPIDGNOLVNPETNSIVSGSTGTTKPKPGIPIVNGGVPVDEEAKDQ 461
QY 171 GDETMGNDYERAMWTGAMGOITTPVTNLVHGSTA--VGPASETORGTPG----- 219
Db 462 ADK--GKD--GLIVPPTNSINKDPVTNTQYSMTGNINP--ETGKVIPLGSLPGSLNP 514
QY 220 -----GLYGTNTRFAVRPDTGEIYVRHOTLPRDN--WDOCTEFEMAVTN-- 262
Db 515 SFNTPOOTDELGTGKPVDTVIGLPRDSTGELIDPATRLPIGSVAGDEILTEVLNITDE 574
QY 263 ----VDVQ--PSTENEGLOS-----INP-----NAATG-----ERRVLT 290
Db 575 VTGLPIIDLETGLPRDPVSGILPOLPNCITLVDPSNKKRIPGSHSGFINGISGEOSHEKDPST 634
QY 291 GVPCKTGT--MMOPD-----AETGEFLMADTNR-----ONMIES 323
Db 635 GKPLDNTGTHPFDEDESGSLINPETGDKLQSGSHSGTFMPVPGKPOGENGIMTPEQILEA 694
QY 324 IDENGIVTYNE-----DAI-----LKELDVERD-----VCPTEL 352
Db 695 L--NKLPTSNENVISPRSSDAVDPDPTNTWMNKISGQTYQVDGKTI PGSASAVIHTAL 752
QY 353 G--GRDMPASAL--NPDGSIYFIPLANNICYDMAVDO-----EFTSMDYNTSNVTKLPP 403
Db 753 GTPOTDPTTGLPSDPSTGLPIGPNVLPDQTEQIKGSVPYSLVKEKNVTEAAY 812
QY 404 GKDMIGRIDAIDISTGRILMSVERAANYSPLSTG-----GGVLFNNGGTDRIYFRA-- 454
Db 813 GL-----PVDPKTGFIDPI-----SYLPEAKNGELIDPISGKIYSGSIAGFISGKAG 860
QY 455 ----LSOETGETLMOT-----RLATVASGOAISYEVDGMQYVAIAG-----GGV 494
Db 861 SOSKSSDESNGPIDPSTNMPYDPKGLIDPESGIAIDNSVSGV-FATVPGTAAPKKGCV 919

```

Search completed: May 24, 2002, 10:20:08
 Job time: 321 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2002, 10:21:03; Search time 54.76 Seconds
(without alignments)
880.877 Million cell updates/sec

Title: WALICK-934-125.PCP
Perfect score: 2689
Sequence: 1 MKPTSLIMASAGALLALIAAP.....GMQYVAINGGVSYGSGINS 502

Scoring table: BIOSUM62
Gapop 10.0, Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: PIR-71:1:
1: p1r1:
2: p1r2:
3: p1r3:
4: p1r4:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	487	18.1	742	2 A49340	alcohol dehydrogenase
2	481	17.9	601	2 B95863	alcohol dehydrogenase
3	479.5	17.8	708	2 S52317	alcohol dehydrogenase
4	475	17.7	742	2 J50326	alcohol dehydrogenase
5	474.5	17.6	738	2 S14270	alcohol dehydrogenase
6	435.5	16.2	626	2 J00706	alcohol dehydrogenase
7	413	15.4	623	2 B83399	alcohol dehydrogenase
8	391.5	14.6	573	2 S68591	alcohol dehydrogenase
9	339.5	12.6	796	1 JY0107	alcohol dehydrogenase
10	339.5	12.6	796	2 H90644	alcohol dehydrogenase
11	339.5	12.6	796	2 H85495	alcohol dehydrogenase
12	331.5	12.3	796	2 AG0523	alcohol dehydrogenase
13	315	11.7	801	1 S00943	alcohol dehydrogenase
14	287.5	10.7	808	1 OPRKX	alcohol dehydrogenase
15	271	10.1	809	2 B98314	alcohol dehydrogenase
16	271	10.1	809	2 A12968	alcohol dehydrogenase
17	267	9.9	803	2 F83360	alcohol dehydrogenase
18	258	9.6	778	2 G98821	alcohol dehydrogenase
19	258	9.6	778	2 A13064	alcohol dehydrogenase
20	250.5	9.3	809	2 A55547	alcohol dehydrogenase
21	230	8.6	639	2 JC4881	alcohol dehydrogenase
22	209.5	7.8	524	2 A41378	alcohol dehydrogenase
23	153.5	5.2	407	2 A82580	alcohol dehydrogenase
24	140.5	5.1	668	2 C75264	alcohol dehydrogenase
25	137.5	5.1	839	2 D97013	alcohol dehydrogenase
26	137	5.0	827	2 F64512	alcohol dehydrogenase
27	134.5	4.9	866	2 A82284	alcohol dehydrogenase
28	132.5	4.9	392	2 F91050	alcohol dehydrogenase
29	130.5	4.9	392	2 F91050	alcohol dehydrogenase

30	130.5	4.9	392	2 B85895	probable dehydrogenase
31	130.5	4.9	392	2 AG0821	probable dehydrogenase
32	129.5	4.8	392	2 G65027	probable dehydrogenase
33	129.5	4.7	525	2 A99953	probable dehydrogenase
34	125.5	4.7	593	2 S00128	probable dehydrogenase
35	125	4.6	407	2 C82804	probable dehydrogenase
36	125	4.6	2535	2 AC0304	probable dehydrogenase
37	124	4.6	3972	2 S75251	probable dehydrogenase
38	121.5	4.5	943	2 JC4081	probable dehydrogenase
39	120	4.5	1354	2 AG0538	probable dehydrogenase
40	118.5	4.4	799	2 T48889	probable dehydrogenase
41	116.5	4.3	424	2 T29127	probable dehydrogenase
42	116.5	4.3	1350	2 AF2005	probable dehydrogenase
43	116	4.3	1156	2 C87371	probable dehydrogenase
44	116	4.3	1441	2 B86807	probable dehydrogenase
45	115.5	4.3	1197	2 D82696	probable dehydrogenase

ALIGNMENTS

RESULT 1
A49340 alcohol dehydrogenase (EC 1.1.1.-) precursor - Acetobacter pasteurianus (strain NC113)
C:Species: Acetobacter pasteurianus
C:Date: 07-Apr-1994 #sequence_revision 07-Apr-1994 #text_change 20-Jun-2000
C:Accession: A49340
R:Takekura, H.; Kondo, K.; Horinouchi, S.; Beppu, T.
J. Bacteriol. 175, 6857-6866, 1993
A:Title: Induction by ethanol of alcohol dehydrogenase activity in Acetobacter pasteurianus
A:Reference number: A49340; MUID:94042848
A:Accession: A49340
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-742 <YAK>
A:Cross-references: GB:D13893; NID:g517067; PIDN:BAA40252.1; PID:g452586
C:Superfamily: glucose dehydrogenase (pyridoxal-5-phosphate-quinone-dependent oxidoreductase)
C:Keywords: alcohol metabolism; NAD; oxidoreductase

Query Match	18.1%;	Score 487;	DB 2;	Length 742;
Best Local Similarity	26.5%;	Pred. No. 3.5e-27;		
Matches 161;	Conservative 81;	Mismatches 215;	Indels 150;	Gaps 21;
QY	11	AGALALAPAFVAVPVDELL--ANPAGWISGQNGEYRSHSPLOITENVQLO 68		
DB	23	AAALPYAAVPAADGGGCTGEALIHADHPENNLVSGRTYSQRYSPLOQINRSVGDIA 82		
QY	69	LVARAGMGPCKVQ--VPLIHGVMYLANPGDYQALDAKTDGLWEHRQLP-NIA----- 122		
DB	83	LAWYTTLDITNGEATPLVVDGIMVATTNMSKWEALDAATGKILMOYDKVPGNTADKGC 142		
QY	123	--TIN-----I 126		
DB	143	CDTVNRGAGYWNKGFVWGTEDGRLVAADAKTGKWEVNTIPADSLGKQSYVDGAVR 202		
QY	127	VANGITVAGSTQCYSPFC--FVSGHDSATGEELWRYFIPRAGEGD----- 172		
DB	203	VAKGVLVIGN--GGSEFARGFVSAFPAETGKIKRFRVTPNNKNRBPDAVADNVIMSKA 260		
QY	173	-ETWGNDEYARMT-----GAMGQITDPTNIVHYSTAVGPASFTORTPGTLY-- 223		
DB	261	YKTWGP--KGAVRGGGGGTWDSLVYDPSDLVY--LAVG-----NSPNVYKRSR 309		
QY	224	--GTN-----TRFVRPVDGELVWRHQTLPDRNWDCECTFEMATNTVDVOSTEBEGLOST 277		
DB	310	GIGSNLEFLGSIYALKRFEIGEVYVHMQANPMQMDYTSVOQITLITLDMVY----- 357		
QY	278	NPAATGE--RRLVGVPCKTGTMOQDAETGEFLMARPTVONNNIESIDE--NGIVTVNED 335		
DB	358	-----NGEMRHVIMHAP--KNGFFVYLDATGTEFLGKRVYQNNANGIDPLTGRPIYNDP 411		
QY	336	ALLKEIDVEIDVCPFLIGRDWPSALNPDGSIYPIPLNNVY-----DMAA 382		

Db 412 GLVTLGKFWYIGPGLAHNMGMANVSPKTHLYLPAHOLPEFGKNOVGCFKPHDAMN 471
 QY 383 VDOEFISMVYVTSNTVTLPPCKDMIGRIDAIDISTGRTLSVYRAANSPVLSTGGV 442
 Db 472 VGLDMKGLPPTPE-ARTAVIKIDLHGLMLAMPVKKETVWKIDHGMNGVLATGSDL 530
 QY 443 LFNGGTDYFRFALSGESETLMQTRLATVAGSQAISYEDVGMQYVA-----IAGG 492
 Db 531 LFGGLANEFHAYDATNMGSDLYKFPDAQSGIASPMTYSVNGKYAVEVGMGGIPIISMG 590
 QY 493 GVSYSGG 499
 Db 591 GVGRTSG 597

RESULT 2
 E95863
 alcohol dehydrogenase (acceptor) (EC 1.1.99.8) large subunit [imported] - *Sinorhizobium*
 C:Species: *Sinorhizobium meliloti*
 C>Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 14-Sep-2001
 C:Accession: E95863
 R:Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Hernan
 Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001
 A:Title: The complete sequence of the 1,683-kb pSymb megaplasmid from the N2-fixing endo
 A:Reference number: A95842; MUID:21396508; PMID:11481431
 A:Accession: E95863
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-601 <KUP>
 A:Cross-references: GB:AL591985; PIDN:CA048573.1; PID:G15140045; GSPDB:GM00167
 A:Experimental source: strain 1021, megaplasmid pSymb
 R:Galibert, F.; Finan, T.M.; Long, S.R.; Publer, A.; Abola, P.; Ampe, F.; Barloy-Hubler,
 L.; Hymn, R.W.; Jones, T.
 Science 293, 668-672, 2001
 A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,
 hebaunt, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.
 A:Title: The composite genome of the legume symbiont *Sinorhizobium meliloti*.
 A:Reference number: A96039; MUID:21368234; PMID:11474104
 A:Contents: annotation
 C:Genetics:
 A:Gene: SMD20173
 A:Genome: plasmid
 C:Keywords: oxidoreductase

Query Match 17.9%; Score 481; DB 2; Length 601;
 Best Local Similarity 25.8%; Pred. No. 6,9e-27;
 Matches 155; Conservative 90; Mismatches 206; Indels 150; Gaps 22;
 QY 14 LALLAFAFAQVTPVDEL--LANPAGEMISYQNOENYRSPLOTOTENNQOLUW 71
 Db 8 LAINISGGGAQVAFANDELQKLIDP-NQWAIQTGDYANLKYSKLDQINKDNVGLAYW 66
 QY 72 --ARGMQPKVQVYPLIHGVNVLNP-----GDV----- 99
 Db 67 TFGTVLRBH-EGSPVIGDLMVHTPPNVTVALDLKSDQIWKYEPKODPNVPIVK 125
 QY 100 -----IQAIDAKTGDLIW-----EHRQLPNATLNIYANGV 131
 Db 126 CPTNAGVAYADNKTIFLHQADPTVALDAKGTGYINSVKMGDATKGTNTATVMPVKDI 185
 QY 132 IVAGSTCOVSPGCFVSGHDSATGEELRNFT-----PRAGEGDE 173
 Db 186 LVGISGGERGVAG-HVTAIYSMADKVIAMRGYSKGPDSIDLIDPEKTHLGKPVGDSGLT 244
 QY 174 TWGNDYEARMTG--AMGQITVDPVNLVHGSTAVGPAASEQKTPGGLTYGNTFRA 230
 Db 245 TWEQD---QMKIGGTTMGWSTYDPEENLIVYGTGNPSTWNPOR--PEDNRM-SMTIFA 298
 QY 231 VRPDTGELVWHQTLPRDNNDQECTEMAVTNDVQPTSTEMGLDSINPNATGRRVLT 290

Db 299 RDVDTGAKMLYQMTPDHEDWDYGVNEMITLREQIDG-----DRKILT 342
 QY 291 GVPCKTGMQFPAEGERL-----WADTNQNMIE-----SIDENGI 329
 Db 343 HFD-RNGFGTYMDRYGVELVAEKYDPTVNMATEVMDPKDAGRPOVAQSTONG- 400
 QY 330 VTVNEDAIKEIDVEYDVCPPFLGGRDMPASALPDGSIYFIPLANNVCYMAVDOEFTS 389
 Db 401 ----EDTNT-----GCPALGTRKDDQPAVSPKTELFYVPTNHYCMDEPFRVSYTA 450
 QY 390 MDVYNTSNTVTLPPCKDM--TGRIDAIDISTGRTLSVYRAANSPVLSTGGVLFNG 446
 Db 451 GQPIYGATLSMTP-KDSHCGMGNFIAMDNKEGKIKSLPEPFSVSGALATRAQVDFG 509
 QY 447 GDRYFRALSGESETLMQTRLATVAGSQAISYEDVGMQYVAIAGS-----GVSYGSLN 501
 Db 510 TLEGITAKVADATGKELREKTPSGVIGNVYARBGKYAVAVLGVGMAGIGLAQLT 569
 QY 502 S 502
 Db 570 N 570

RESULT 3
 S52317
 quinohemoprotein ethanol dehydrogenase (EC 1.1.99.-) type 1 precursor - *Comamonas tes*
 C:Species: *Comamonas testosteroni*
 C>Date: 08-May-1995 #sequence_revision 21-Jul-1995 #text_change 02-Jun-2000
 C:Accession: S62366; S62373; S65908; S52317
 R:Stoorvogel, J.; Kraayveld, D.E.; Van Suijs, C.A.; Jongejan, J.A.; de Vries, S.; Dui
 Eur. J. Biochem. 235, 690-699, 1996
 A:Title: Characterization of the gene encoding quinohemoprotein ethanol dehydrogenase
 A:Reference number: S62366; MUID:96184549
 A:Accession: S62366
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-708 <STOI>
 A:Cross-references: EMBL:X81880; NID:9663195; PIDN:CAA57464.1; PID:9663196
 A:Accession: S62373
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 32-54 <STO2>
 R:de Jong, G.A.H.; Geerloff, A.; Stoorvogel, J.; Jongejan, J.A.; de Vries, S.; Duine,
 Eur. J. Biochem. 230, 899-905, 1995
 A:Title: Quinohemoprotein ethanol dehydrogenase from *Comamonas testosteroni*. Purific
 A:Reference number: S65908; MUID:95324580
 A:Accession: S65908
 A:Molecule type: protein
 A:Residues: 32-50, 'X', 52-54, 477-483, 'X', 485-490 <DED>
 A:Experimental source: ATCC 15667
 C:Genetics:
 A:Gene: qhdh
 C:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)
 C:Keywords: oxidoreductase; quinoprotein
 P:1-31/Domain: signal sequence #status predicted <SIG>
 F:32-708/Product: quinohemoprotein ethanol dehydrogenase type 1 #status experimental

Query Match 17.8%; Score 479.5; DB 2; Length 708;
 Best Local Similarity 25.0%; Pred. No. 1.1e-26;
 Matches 148; Conservative 84; Mismatches 228; Indels 121; Gaps 18;
 QY 3 PISILMASAGALALLAFAFAQVTPV-----DELANP-AGMISYQNOEN 50
 Db 12 PGRWVWLLAACG--SAAAFQOTGPAQAQAAAVQVDDDFIRNAAATPMPITGVDAE 69
 QY 51 YRSPILTQTTNNGOOLVARGMQPK-VQVPLIHGVNVLANGVDYIAIDAKTD 109
 Db 70 TRYSRLDQINANVNDLISWSYNLESTRGVEATFPVVDGIMTVASWSVVAIDRTGN 129
 QY 110 LIMEHRRLPNT--ALNIVANG----- 130
 Db 130 RIWYTDPOIDNSTGFGCCDVVNRGVALLMRKGYVVGAMDCRLIALDAATGKEVWHQNTPE 189


```

QY 173 -----ETWGNDEYARMWMTGA-----WGLTYDPVNLVHYGSTAVGPASETOR 215
      |||
Db 255 VLMNRKAYOTMS-----PTGAMTROGGGGTWMSIYDDPVALYLTVGVNGSPMNYKKR 307
      |||
QY 216 GTPGGTLYGTNRRAVRDPDTEIWRHQTLPDNNDOCECEEMAVTVNDQPTBMEGIC 275
      |||
Db 308 SEKGNGNLFGLGIVALKPDETEYWHFQETPMQDMFTSQOIMTIDLPI----- 357
      |||
QY 276 SINPNAAGERRVLTGVCCTGTMMQPDATGEFLNARDNTYONMIESID-ENGITVYNE 334
      |||
Db 358 -----NGETRHYIVAHKNGFFTYITDGTGETISGKNVYVMAASGIDPKTGRIYNP 410
      |||
QY 335 DALKELDEYDVCPTFLGGRWPSAALNPDGTYEPLNNVCYDMAAIDQDEFT-SMVOY 393
      |||
Db 411 DALTYLGGKEWYGIPIGLGSHNFAAMAFSPKGLVIYIPAOQVPELYTNQVGGFTPHPSW 470
      |||
QY 394 NTS---NVTKLPPG-----KDMIGRIDAIDISTGRTLSVERAANYSPVLTSGGV 442
      |||
Db 471 NLGLDMKNGVPIPSPEAKQAFYKDLGKMTVAMDPOKQAEAMRVDRHKGPMNGGILATGSDL 530
      |||
QY 443 LFNCGTDRYFRALSGETGELTMQTRLATVVASGQATSEYEDQMQYVAIAG 492
      |||
Db 531 LFGSLANGFERHAYDATNGSDLFHFAADSGIAPVPTILANGKQYAAVEVG 580
      |||

```

RESULT 6

```

JQ0706
alcohol dehydrogenase (acceptor) (EC 1.1.99.8) alpha chain precursor - Methylobacterium
N:Alternate names: methanol dehydrogenase 62k large chain
C:Species: Methylobacterium extorquens
C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 08-Oct-1999
R:Accession: JQ0706; 507908
R:Author: D.J. Morris, C.J. Nunn, D.N.; Anthony, C.; Lidstrom, M.E.
A>Title: Nucleotide sequence of the Methylobacterium extorquens AM1 mofx and mofx genes
A:Reference number: JQ0706; MUID:90337342
A:Accession: JQ0706
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-626 <AND>
A:Cross-references: GB:W31108; NID:9150017; PIDN:AA25380.1; PID:9150018
R:Nunn, D.N.; Day, D.; Anthony, C.
Biochem. J. 260, 857-862, 1989
A>Title: The second subunit of methanol dehydrogenase of Methylobacterium extorquens AM1
A:Reference number: S04644; MUID:89350892
A:Accession: S07908
A:Molecule type: protein
A:Residues: 28-50, XX, 53 <NUN>
A:Note: the source is designated as Methylobacterium extorquens AM1
C:Comment: This enzyme oxidizes methanol to formaldehyde.
C:Genetics: mofx
A:Gene: mofx
C:Keywords: alcohol metabolism; oxidoreductase
F:1-27/Domain: signal sequence #status predicted <SIG>
F:28-626/Product: alcohol dehydrogenase (acceptor) #status experimental <MAT>

```

Query Match

Best Local Similarity 16.2%; Score 435.5; DB 2; Length 626;
Matches 148; Conservative 91; Mismatches 208; Indels 155; Gaps 23;

```

QY 10 SAGALALIA-AP-AFAQVTPYTDLANPAGW-WISYQONQENYRHSPLTOITENWQ 66
      |||
Db 7 SVSALAMIALAPALSSGAVANDKLVELSKSDNVMVPGKNDNSNNFSLDKINKGNKQ 66
      |||
QY 67 LQLVW-ARGMQPGKVOTPLIHGVVY-----LANGVY----- 99
      |||
Db 67 LRPMTFTSTGLNGH-EGAPLVYDGMKTYITSPNNFALGLDDGTILMDKPKQNDPA 125
      |||
QY 100 -----IOAIDAKTGDLIWEHRQLPIATLN 125
      |||
Db 126 RAVACCLVNRGLAVWPGDKPTALITKTOLDGNAVAAALNAEIGETVWK----- 173
      |||

```

```

QY 126 IVANGVIVASTQYSPF-----GC-----FVSHDSATGEELMRNFI----- 164
      |||
Db 174 -VENSIDIKVSTLTITAPYVKKRVIIIGSSGELGVRLTAVDKTGEQVVRVAVGPK 232
      |||
QY 165 -----PRAGEE--GDETWNDEYARMWMTGA-----WGLTYDPVNLVHYGSTAV 207
      |||
Db 233 DILLASDFIKNDPHGOKGLGTGTWEGD--AMKIGGTNGKVAIDPQNTLLYFETGNP 289
      |||
QY 208 GPASETQRTPGCTLYGNTFRAPVRDPNGELVWHRQTLPDNNDOCECEEMAVTVNDYQ 267
      |||
Db 290 APWNETMR--PDNDNR-TWTFEGRADDTGEAKFEYQKTPHDEWD-----YAGVNVMA 338
      |||
QY 268 STEMEGLOSINPNAATGERRVLTGVPCKTGTMMQPDATGEFLNARDNTYONMIESI 324
      |||
Db 339 LSEQKD-----KDGKARLLTHPRNGLIYTLDRDICALYANKLDDT--VNVFRSV 388
      |||
QY 325 DENGIVTNEADALKELD-VEYDVCPTFLGGRWPSAALNPDGTYEPLNNVCYDMAV 383
      |||
Db 389 DLKTGQVPRDPEYGTNRDHLAKDICPSAMGYHNGHSDYDPRELFEFMGINHICMDEPF 448
      |||
QY 384 DQETSMQVYNTSVNKLPPGK-----DMIGRIDAIDISTGRTLSVERAANYSPVLTSGGV 437
      |||
Db 449 MLPYRAGQFVAGVILNMYPGPKGRQNGYEGIKAYNAITDQYKEMERFAVWGVTMA 508
      |||
QY 438 TGGVLFNGGTDYFRALSGETGELTMQTRLATVVASGQATSEYEDQMQYVAIAGVSYG 497
      |||
Db 509 TAGDLVYGYGLDYLTKARSDTDGLMKFKIPSGAIGVPMYTHKGQYVAI-----YTG 563
      |||
QY 498 SG 499
      |||
Db 564 VG 565
      |||

```

RESULT 7

```

B83399
guinoprotein alcohol dehydrogenase PA1982 [imported] - Pseudomonas aeruginosa (strain
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
R:Stover, C.K.; Pham, X.O.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.;
  Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A>Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pa
A:Reference number: A82950; MUID:20437337
A:Accession: B83399
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-623 <STO>
A:Cross-references: GB:AE004624; GB:AE004091; NID:99947973; PIDN:AA05370.1; GSPDB:GN
A:Experimental source: strain PA01
C:Genetics:
A:Gene: exaA; PA1982

```

Query Match

Best Local Similarity 15.4%; Score 413; DB 2; Length 623;
Matches 140; Conservative 93; Mismatches 245; Indels 122; Gaps 20;

```

QY 3 PISILMASAGALAL---LAAPFAQVTPYTDLANP-PPAGEMISYQONQENYRHSPLT 57
      |||
Db 9 PAGLLRPSLHCLAFVAVAGSAGALAKVWEDJANDKTTGDLVIGMGTHQNRSPK 68
      |||
QY 58 QITTEVNGOLQLVWAGM---QPGKVOTPLIHGVVYLANGVYQAIQAIADAKTGDLIWE 113
      |||
Db 69 OVNAQNVKRLTPAMVSFSGDEKQGG--ESQAIYSDGVITYTASYSFLFADAKTKRLMT 127
      |||
QY 114 HRRQLPN-----IATL-NIYA-----NGVIV-----AGST 137
      |||
Db 128 YNHRLLPDDIRCCQVNVNRGAALTYGDKVFETGLDASVYVALNKNKTKGVWKKKPFADGAGYT 187
      |||
QY 138 QOYSP-----FGCF--VSGHDSATGEELMRNFI----- 166
      |||

```

[illegible][illegible]

RESULT 9

JV0107
glucose dehydrogenase (pyrroloquinoline-gulonone) (EC 1.1.99.17) - Escherichia coli

C:Species: Escherichia coli
C:Date: 30-Jun-1992 #sequence_revision 12-Dec-1997 #text_change 16-Jun-2000
C:Accession: D64735; JV0107; A45997; S45201; I41228
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.;
..; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID:97426617
A:Accession: D64735
A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA
A:Residues: 1-796 <BLAT>
A:Cross-references: GB:A000122; GB:U00096; NID:91786315; PIDN:AAC73235.1; PID:917863
A:Experimental source: strain K-12, substrain MG1655
R:Clifton-Jansen, A.W.; Goosen, N.; Fayet, O.; van de Putte, P.
J. Bacteriol. 172, 6308-6315, 1990
A:Title: Cloning, mapping, and sequencing of the gene encoding Escherichia coli quino
A:Reference number: JV0107; MUID:91035240
A:Accession: JV0107
A:Molecule type: DNA
A:Residues: 1-58, 'L', 60-148, 'H', 150, 'KRRCHT', 157-192, 'K', 194-796 <CLE>
A:Cross-references: GB:X51323; NID:941553; PIDN:CAA35706.1; PID:941554
A:Experimental source: strain K12
R:Yamada, M.; Sumi, K.; Matsushita, K.; Adachi, O.; Yamada, Y.
J. Biol. Chem. 268, 12812-12817, 1993
A:Title: Topological analysis of quinoprotein glucose dehydrogenase in Escherichia co
A:Reference number: A45997; MUID:93286127
A:Accession: A45997
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-20 <YAM>
R:Fujita, N.
submitted to the EMBL Data Library, January 1994
A:Reference number: S45181
A:Accession: S45201
A:Molecule type: DNA
A:Residues: 1-796 <FUB>
A:Cross-references: EMBL:D26562; NID:9473770; PIDN:BA005580.1; PID:9473791
A:Experimental source: strain K-12 substrain W3110
R:Yamada, M.; Asaoka, S.; Saler, M.H.; Yamada, Y.
J. Bacteriol. 175, 566-571, 1993
A:Title: Characterization of the gcd gene from Escherichia coli K-12 W3110 and regula
A:Reference number: I41228; MUID:93123180
A:Accession: I41228
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-148, 'H', 150, 'KRRCHT', 157-192, 'K', 194-665, 'H', 667-796 <RES>
A:Cross-references: GB:D12651; NID:9216555; PIDN:BA002174.1; PID:9216556
C:Genetics:
A:Gene: gcd

A:Map position: 3 min

C:Function:

A:Description: catalyzes the oxidation of D-glucose to D-gluconic acid by ubiquinone

A:Pathway: respiratory chain

C:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)

C:Keywords: oxidoreductase, pyrroloquinoline-quinone, respiratory chain, transmembrane

F:11-37/Domain: transmembrane #status predicted <TM1>

F:41-59/Domain: transmembrane #status predicted <TM2>

F:63-81/Domain: transmembrane #status predicted <TM3>

F:96-110/Domain: transmembrane #status predicted <TM4>

F:120-140/Domain: transmembrane #status predicted <TM5>

F:93/95/Binding site: ubiquinone (Arg, Asp) #status predicted

F:466/Active site: Asp #status predicted

Query Match 12.6%; Score 339.5; DB 1; Length 796;

Best Local Similarity 22.5%; Pred. No. 1.7e-16;

Matches 158; Conservative 75; Mismatches 203; Indels 267; Gaps 30;

```

OY 6 LMAS-----AGALLAARFAQVTPYDELANPAGEMWISYQONENRHSPLTQ 58
DB 135 LTMAGFNDPQETNGTSLADATPAEA-ISPVDQ-----DWPAYGRNCGQRFSPKQ 185
OY 59 ITTENGOLOLYMA-----RGMOPKV--QVPLIHGVMYLANPGDVIQAIADAKTD 109
DB 186 INADNVHNKEMAVFRTGDKQPNPDEITNEVPIKVGDTLYLCTAHQRLFALDASGR 245
OY 110 LIMEHRRL-----RGMOPKV--QVPLIHGVMYLANPGDVIQAIADAKTD 109
DB 246 EKMVHDPPELKTNEFQHTGCVSYHEAKETAPEVMADCPRIILPVNDRLIAINAE 305
OY 119 -----PNTATLIYAN-----GVVAGS--TCQYS--PGCFV 147
DB 306 NCKLCEETANKGVNLQSNMDETRKPGLEYEPTSPRIITDKIYVAGSVTNESTRETSVI 365
OY 148 SGHSATGEELMRYNYFIRA-----GEEDETKNDYDEARMGAMQIYDPTVLY 200
DB 366 RGFVNTGELLMA--FDPGAKDPALIPSDHTFTNS-----PNSWAPAYDAKLDLY 416
OY 201 HYGSTAVGPASETQRTPGSTLGTMTTRAVRPDGEIYWRHQLPRNMDOECTFEKAV 260
DB 417 YLPGVVTPPIWGNKRTPEQERYASSI-LALNATGKLAWSYQVTHDLMMD-----468
OY 261 TNDVOSTEMEGLOSINPAATGERRVLTGVCKTGTMQDFAETGEL-----310
DB 469 --LPAQPT-----LADITYN--GQKVPIYAPAKTGIFVLDNRNGELVYPAPEKVPQ 518
OY 311 WARDTNYQNMIESIDENGI-----VYVEDAILKEI--DVEYD-----346
DB 519 GAAKGDVTPPTQPFSELSFRPTKDLGADMWGATMFQDLYCRVAFHQRREGIFTPPSQ 578
OY 347 ---VCPTEFLGGRDWPSSALNP-----D 365
DB 579 GTLVFPGNIGMEWGISVDNREVAIANPALPFSKLIIPRGPNMEQPKDAKTGTE 638
OY 366 SGI-----YFIPLNNCYDMAVDOEFTSMVYNTSKLPKPKMIGRIDAIDIST 418
DB 639 SGIOQOYGVPIGVTLN-----PFLS-----PFGIPLCKQAPMGYISALDKT 679
OY 419 GRTLSVERAANYS-----PV-----LSTGGVLPFGCGT-DRYFALSQ 458
DB 680 NEVWVKRRTIGTPDQSMFPMPVPVPEFMGMPLGPISTAGVNLFLAATADYLAAYNMS 739
OY 459 TGETLMOTRLATVASGA--ISYEVDMQYVAI-AGGVSYGS 498
DB 740 NGEKIMOGRLP--AGGQATPTMYEVNGKYVVISAGHGSFCT 780

```

RESULT 10

H90644

glucose dehydrogenase [imported] - Escherichia coli (strain O157:H7, substrain RMD 0509)

C:Species: Escherichia coli

C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Aug-2001

C:Accession: H90644

R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.

gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shingawa, H.

DNA Res. 8, 11-22, 2001

A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and 9

A:Reference number: A59629; MUID:21156231; PMID:11258796

A:Accession: H90644

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-796 <HAV>

A:Cross-references: GB:BA000007; PIDN:BA033551.1; PID:g13339584; GSPDB:GN00154

A:Experimental source: strain O157:H7, substrain RMD 050952

C:Genetics:

A:Gene: EC0128

C:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)

Query Match 12.6%; Score 339.5; DB 2; Length 796;

Best Local Similarity 22.5%; Pred. No. 1.7e-16;

Matches 158; Conservative 75; Mismatches 203; Indels 267; Gaps 30;

```

OY 6 LMAS-----AGALLAARFAQVTPYDELANPAGEMWISYQONENRHSPLTQ 58
DB 135 LTMAGFNDPQETNGTSLADATPAEA-ISPVDQ-----DWPAYGRNCGQRFSPKQ 185
OY 59 ITTENGOLOLYMA-----RGMOPKV--QVPLIHGVMYLANPGDVIQAIADAKTD 109
DB 186 INADNVHNKEMAVFRTGDKQPNPDEITNEVPIKVGDTLYLCTAHQRLFALDASGR 245
OY 110 LIMEHRRL-----RGMOPKV--QVPLIHGVMYLANPGDVIQAIADAKTD 109
DB 246 EKMVHDPPELKTNEFQHTGCVSYHEAKETAPEVMADCPRIILPVNDRLIAINAE 305
OY 119 -----PNTATLIYAN-----GVVAGS--TCQYS--PGCFV 147
DB 306 NCKLCEETANKGVNLQSNMDETRKPGLEYEPTSPRIITDKIYVAGSVTNESTRETSVI 365
OY 148 SGHSATGEELMRYNYFIRA-----GEEDETKNDYDEARMGAMQIYDPTVLY 200
DB 366 RGFVNTGELLMA--FDPGAKDPALIPSDHTFTNS-----PNSWAPAYDAKLDLY 416
OY 201 HYGSTAVGPASETQRTPGSTLGTMTTRAVRPDGEIYWRHQLPRNMDOECTFEKAV 260
DB 417 YLPGVVTPPIWGNKRTPEQERYASSI-LALNATGKLAWSYQVTHDLMMD-----468
OY 261 TNDVOSTEMEGLOSINPAATGERRVLTGVCKTGTMQDFAETGEL-----310
DB 469 --LPAQPT-----LADITYN--GQKVPIYAPAKTGIFVLDNRNGELVYPAPEKVPQ 518
OY 311 WARDTNYQNMIESIDENGI-----VYVEDAILKEI--DVEYD-----346
DB 519 GAAKGDVTPPTQPFSELSFRPTKDLGADMWGATMFQDLYCRVAFHQRREGIFTPPSQ 578
OY 347 ---VCPTEFLGGRDWPSSALNP-----D 365
DB 579 GTLVFPGNIGMEWGISVDNREVAIANPALPFSKLIIPRGPNMEQPKDAKTGTE 638
OY 366 SGI-----YFIPLNNCYDMAVDOEFTSMVYNTSKLPKPKMIGRIDAIDIST 418
DB 639 SGIOQOYGVPIGVTLN-----PFLS-----PFGIPLCKQAPMGYISALDKT 679
OY 419 GRTLSVERAANYS-----PV-----LSTGGVLPFGCGT-DRYFALSQ 458
DB 680 NEVWVKRRTIGTPDQSMFPMPVPVPEFMGMPLGPISTAGVNLFLAATADYLAAYNMS 739
OY 459 TGETLMOTRLATVASGA--ISYEVDMQYVAI-AGGVSYGS 498
DB 740 NGEKIMOGRLP--AGGQATPTMYEVNGKYVVISAGHGSFCT 780

```

RESULT 11

H85495

glucose dehydrogenase [imported] - Escherichia coli (strain O157:H7, substrain EDL933)

C:Species: *Escherichia coli*
 C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
 C:Accession: H85495
 R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
 M.; Miller, L.; Golubek, E.J.; Davis, N.W.; Lim, A.; Dimantanta, E.; Potamoukis, K.; Apodaca,
 Nature 409, 529-533, 2001
 A:Title: Genome sequence of enterohemorrhagic *Escherichia coli* O157:H7.
 A:Reference number: AB5480; MUID:21074935; PMID:11206531
 A:Reference: H85495
 A:Accession: H85495
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-796 <STO>
 A:Cross-references: GB:AE005174; MTD:q12512839; PIDN:AG54428.1; GSPDB:GN00145; UMGSP:201
 A:Experimental source: strain 0157:H7, substrain EDL933
 C:Genetics:
 A:Gene: gcd
 C:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)

Query Match 12.6%; Score 339.5; DB 2; Length 796;
 Best Local Similarity 22.5%; Pred. No. 1.7e-16;
 Matches 158; Conservative 75; Mismatches 203; Indels 267; Gaps 30;

QY 6 LIMS-----AGALILAPAFQVPTVDLLANPPAGEMISGQNGENYRHSPLTQ 58
 DB 135 LTMAGFNDPQELNGTLRADATPA-ATSSIADE-----DWAYGRNREGORYSPK 185
 QY 59 ITTENVGOLQVMA-----KGMQPKV--OYPLHDGVMYLANPGDVLQADAKTGD 109
 DB 186 ITADNVHNLKEAWFRTGDKQPNDEITNEVPIKQVDTLYCTAHQRLFALDAASGK 245
 QY 110 LIMEHRQL----- 118
 DB 246 EKMHPDLQTDSSFOHYTCRGVSYHEAKADTASPEVIADCPRIILFVNDGRLEAVNAE 305
 QY 119 -----PNATLNIYAN-----GVIVAGS--TCQYS--PFGCFV 147
 DB 306 NGKLCETFAKNGVNLQSNMPTKGLYEPSPITTDKTIYVAGSYTNDSESTRETSGLV 365
 QY 148 SGHDSATGEELMRYFTIPRAAGEBDGTDGNDYBARMT--GAMGQITVDPTNLYHYS 205
 DB 366 RGFVDVNSGKLMA--FDE--GAKDPTIPADDEHAFENSPSWAPAAADAKLDLYLPMG 421
 QY 201 HYGSTAVGPASETORGTGGTLYGTNTRPAVRDPTGEIYVHQTLPRDNMOECTFEKMY 265
 DB 417 YLPMGVTPDIMGNRTPEDEYASSI-LALNATGKLAMSVOYVHNDLMDMDLPAQTLADIV 480
 QY 261 TNVVOVSTEMEGLQSNPNAATGERRVLTGVPCKTGTMMQDAETGEF----- 309
 DB 469 --LPAQPL-----LADITVN--GOKVPVLYAPAKTGNIFVLDNRNGELVYAPAPKPVQ 518
 QY 311 WARDNTQNMIESIDENGI-----VTNEDAILKEL--DVEYD----- 346
 DB 519 GAAGKDYVTPQPESELSEFRTKDLSCADWGMATMFDQVCRVMEHQMYEGEITPPSBO 578
 QY 347 ---VCPTFLGGRDMPASALNP----- 365
 DB 579 GLTVFPGNLMGEMWGISVDNREVALANPMALFEVSKLIPRGGNMPEPPKADKGTGE 638
 QY 366 SGI-----YIPLANNCYDMAVAVDOEFTSMOYNTSNYTKLPKGMIGRIDAIDIST 418
 DB 639 SGIOPOYGVPIVGTN-----PFLS-----PFGLPCKOPANGYISALDLTKT 679
 QY 419 GRTLMGVERAANYS-----PV-----LSTGGVLFNGGT--DRYFRALSOE 458
 DB 680 NEVVMKKRIGTPODSMPFMPVYVFPNMGMDLGGPISTAGNVLFIAATADNLTAKYNS 739
 QY 459 TGETLMOTRLATAYASGA--ISYEVDGMOYVAI--AGGVSYS 498
 DB 740 NGKLMQGRLP--AGGQATPTEYVNGKQYVVISAGHGSFGT 780

RESULT 12

AG0523
 glucose dehydrogenase [imported] - *Salmonella enterica* subsp. *enterica* serovar Typhi
 C:Species: *Salmonella enterica* subsp. *enterica* serovar Typhi
 A:Note: this species has also been called *Salmonella typhi*
 C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 27-Nov-2001
 C:Accession: AG0523
 R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Main, J.; Church
 S.; Mout, S.; Gaur, P.
 Nature 413, 848-852, 2001
 A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens,
 A:Title: Complete genome sequence of a multiple drug resistant *Salmonella enterica* se
 A:Reference number: AB0502; PMID:11677608
 A:Accession: AG0523
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-796 <PAR>
 A:Cross-references: GB:AL513382; PIDN:CAD01327.1; PID:q16501455; GSPDB:GN00176
 C:Genetics:
 A:Gene: STD191
 C:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)

Query Match 12.3%; Score 331.5; DB 2; Length 796;
 Best Local Similarity 22.8%; Pred. No. 6.6e-16;
 Matches 157; Conservative 67; Mismatches 227; Indels 237; Gaps 29;

QY 6 LIMS-----AGALILAPAFQVPTVDLLANPPAGEMISGQNGENYRHSPLTQ 58
 DB 135 LTMAGFNDPQELNGTLRADATPA-ATSSIADE-----DWAYGRNREGORYSPK 185
 QY 59 ITTENVGOLQVMA-----KGMQPKV--OYPLHDGVMYLANPGDVLQADAKTGD 109
 DB 186 ITADNVHNLKEAWFRTGDKQPNDEITNEVPIKQVDTLYCTAHQRLFALDAASGK 245
 QY 110 LIMEHRQL----- 118
 DB 246 EKMHPDLQTDSSFOHYTCRGVSYHEAKADTASPEVIADCPRIILFVNDGRLEAVNAE 305
 QY 119 -----PNATLNIYAN-----GVIVAGS--TCQYS--PFGCFV 147
 DB 306 NGKLCETFAKNGVNLQSNMPTKGLYEPSPITTDKTIYVAGSYTNDSESTRETSGLV 365
 QY 148 SGHDSATGEELMRYFTIPRAAGEBDGTDGNDYBARMT--GAMGQITVDPTNLYHYS 205
 DB 366 RGFVDVNSGKLMA--FDE--GAKDPTIPADDEHAFENSPSWAPAAADAKLDLYLPMG 421
 QY 201 HYGSTAVGPASETORGTGGTLYGTNTRPAVRDPTGEIYVHQTLPRDNMOECTFEKMY 265
 DB 417 YLPMGVTPDIMGNRTPEDEYASSI-LALNATGKLAMSVOYVHNDLMDMDLPAQTLADIV 480
 QY 261 TNVVOVSTEMEGLQSNPNAATGERRVLTGVPCKTGTMMQDAETGEF----- 309
 DB 469 --LPAQPL-----LADITVN--GOKVPVLYAPAKTGNIFVLDNRNGELVYAPAPKPVQ 518
 QY 311 WARDNTQNMIESIDENGI-----VTNEDAILKEL--DVEYD----- 346
 DB 519 GAAGKDYVTPQPESELSEFRTKDLSCADWGMATMFDQVCRVMEHQMYEGEITPPSBO 578
 QY 347 ---VCPTFLGGRDMPASALNP----- 365
 DB 579 GLTVFPGNLMGEMWGISVDNREVALANPMALFEVSKLIPRGGNMPEPPKADKGTGE 638
 QY 366 SGI-----YIPLANNCYDMAVAVDOEFTSMOYNTSNYTKLPKGMIGRIDAIDIST 418
 DB 639 SGIOPOYGVPIVGTN-----PFLS-----PFGLPCKOPANGYISALDLTKT 679
 QY 419 GRTLMGVERAANYS-----PV-----LSTGGVLFNGGT--DRYFRALSOE 458
 DB 680 NEVVMKKRIGTPODSMPFMPVYVFPNMGMDLGGPISTAGNVLFIAATADNLTAKYNS 739
 QY 459 TGETLMOTRLATAYASGA--ISYEVDGMOYVAI--AGGVSYS 498
 DB 740 NGKLMQGRLP--AGGQATPTEYVNGKQYVVISAGHGSFGT 780

Db 753 GQATPMTVEVNGKQYVVISAGHGSGFT 780

RESULT 13

S00943

glucose dehydrogenase (pyrroloquinoline-quinone) (EC 1.1.99.17) - Acinetobacter calcoac

C:Species: Acinetobacter calcoaceticus

C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999

C:Accession: S00943

R:Clifton-Jansen, A.M.; Goosen, N.; Odle, G.; van de Putte, P.

Nucleic Acids Res. 16, 6228, 1988

A:Title: Nucleotide sequence of the gene coding for quinoprotein glucose dehydrogenase

A:Reference number: S00943; MID:88289368

A:Accession: S00943

A:Molecule type: DNA

A:Residues: 1-801 <CLE>

A:Cross-references: EMBL:X07235; NID:g38711; PID:CAA30222.1; PID:g38712

A:Experimental source: strain LMD 79.41

C:Function:

A:Description: catalyzes the oxidation of D-glucose to D-gluconic acid by ubiquinone

C:Pathway: respiratory chain

C:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)

C:Keywords: oxidoreductase; pyrroloquinoline quinone; respiratory chain; transmembrane

F:9-35/Domain: transmembrane #status predicted <TM1>

F:39-57/Domain: transmembrane #status predicted <TM2>

F:61-79/Domain: transmembrane #status predicted <TM3>

F:94-108/Domain: transmembrane #status predicted <TM4>

F:118-137/Domain: transmembrane #status predicted <TM5>

F:91,93/Binding site: ubiquinone (Arg, Asp) #status predicted

F:471/Active site: Asp #status predicted

Query Match

Best Local Similarity 11.7%; Score 315; DB 1; Length 801;

Matches 148; Conservative 76; Mismatches 192; Indels 28; Gaps 30;

```

QY 20 PAFQAVTVDLTPAGEMISYXGONENYRHSPLTQTTNVCOLQVARGMOPK 79
Db 152 PETAQAVGVAE-----SDMPATGRTQAGVAKRISPKQINDQVNDLVAWT--LRGD 202
QY 80 V-----QVPLIHGQVWLANPGDVYQAIIDAKTGDLWEHRRQL----- 118
Db 203 LKTNDSETTNOVTPRIKINNMFTCAHQOLIAIDPATGKEKMRDPRKTKDSFOHLT 262
QY 119 ----- 118
Db 263 CRGVYTYDANNTEFATSLQSSKSTOCPRKVFVVDGRLVAVNADTGACDFFGNG 322
QY 119 -----PNTATLNIIVANGVIVAGS--TCQYS---PFGCVFSGHDSAT 154
Db 323 QVNLQEFMPYAPGYNPTSPGIVT---GSTVVIAGSVTDNYSNKEPSG--VINGYIVNT 377
QY 155 GEBIMRNYETPRAGE-----EGDETWNQDYEARMTGAMQOITYDPTNLVHIGSTAVG 208
Db 378 GKLLM--VFDTGADPDNAMPGEETFEVINS-----PAMAPLAVYDKLEIV--YVPTGV- 427
QY 209 PASETGTGTPGGTILYG-----TTFRAVRBPDEGEIYWHQTLPRNMPOECFEEF 258
Db 428 -----GTF--DIMGDRETELKERYANSLAINASTGKLWVNFQTHHDMNDVPSQP 478
QY 259 MYTNVNDVOPSTMEGLQSLNP--NAATGERFVLGVP-----CTKG--TMMQF 302
Db 479 SLADIKRNAGQVPAIYVLTGTGNAFVLDNR--NGQPIYVTEKRPVQGYKRPQTKGEF 536
QY 303 DAETGEF-----LMA-----RDTNTOANNIESIDENGITVNE 334
Db 537 YSKTQPFSDMLAPQDLTKDLMGATMLDLMCRVSEKRLNDGITYTPSENGTL----- 592
QY 335 DALKELDVEYDCPTFLGGRDWPALNDS-----GIYFT----- 371
Db 593 -----VPGNLGVFEWEGMSVNPDRQVAVMNPITGLPEVSRLLPADPNRAQTAK 640
QY 372 -----PLNNVCYMAAVDDEFTSMQVYNTSVNTKLIPGKDMIGRIDAIIDISTGRTL 422

```

Db 641 GAGTEGQVDPMPGVY-----GVEISAFSLPGLPCPKQAKRYVAGVDLKTREVV 690

QY 423 W-----SVERAANYSPLYLSTGGVYLFNFGCT--DRYFALSQETGEL 463

Db 691 WKRIKRTIDSLPNLFQUPAVKIGVPGIGGSISTRAGNMVVGATQDNYLRAFNVTNGKLL 750

QY 464 MQRRLATVASSGA--ISFVDMQVYAL--AGGVSYGS 498

Db 751 WEARL--AGGATPMTYEINGKQYVIMAGHGSGFT 786

RESULT 14

QPKEX

glucose dehydrogenase (pyrroloquinoline-quinone) (EC 1.1.99.17) - Gluconobacter oxyda

C:Species: Gluconobacter oxydans

C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 22-May-1998

C:Accession: S17716; S19265

R:Clifton-Jansen, A.M.; Dekker, S.; van de Putte, P.; Goosen, N.

Mol. Gen. Genet. 229, 206-212, 1991

A:Title: A single amino acid substitution changes the substrate specificity of quinop

A:Reference number: S17716; MID:92017653

A:Accession: S17716

A:Molecule type: DNA

A:Residues: 1-808 <CLE>

A:Cross-references: EMBL:X62710

R:Goosen, N.

submitted to the EMBL Data Library, February 1992

A:Reference number: S19265

A:Accession: S19265

A:Molecule type: DNA

A:Residues: 1-212/A, 214-808 <GOO>

A:Cross-references: EMBL:X62710; NID:g58416; PID:g58417

C:Genetics:

A:Gene: gdh

C:Function:

A:Description: catalyzes the oxidation of D-glucose to D-gluconic acid by ubiquinone

C:Pathway: respiratory chain

C:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)

C:Keywords: oxidoreductase; pyrroloquinoline quinone; respiratory chain; transmembran

F:9-28/Domain: transmembrane #status predicted <TM1>

F:35-54/Domain: transmembrane #status predicted <TM2>

F:60-76/Domain: transmembrane #status predicted <TM3>

F:94-110/Domain: transmembrane #status predicted <TM4>

F:122-138/Domain: transmembrane #status predicted <TM5>

F:91,93/Binding site: ubiquinone (Arg, Asp) #status predicted

F:470/Active site: Asp #status predicted

Query Match

Best Local Similarity 10.7%; Score 287.5; DB 1; Length 808;

Matches 140; Conservative 82; Mismatches 229; Indels 251; Gaps 25;

```

QY 13 ALALAAPAAQVTPYVDELTA-----NPPAGEMISYXGONENYRHSPLTQITT 61
Db 128 AVLALFASLTDPDHDISGELPQIANASPADPDNPASBMAHAGRTQAGDWSPLINDIA 187
QY 62 ENVGQQLQVW-----ARGMOPKAV--QVPLIHGQVWLANPGDVYQAIIDAKTGDLW 112
Db 188 TIVSNLKVAMHHTTKMMNSNDPGQOTNEATPIEFNNLTLYWCSLHOKLFAVDGATGNVKA 247
QY 113 EHRRL--PNTATLNI-----VANGYIV----- 133
Db 248 VYDPKLIQINPFGHLCGVSFHEPRANAMSDGNPAPTDCAKDSILIPVNDGRLVEVDAD 307
QY 134 AGSTC-----QYSEFGCV----- 147
Db 308 TKTSTGCGNNGEIDLKRVNPQPTTPGQYETSPVITDKLIANSALINDNSYKQASCA 367
QY 148 -SGHDSATGEEEL-----RNYFTPRAGEGETGNGNYEARMTGAMQOITYDPTNLVHY 202
Db 368 TQAFVYVYTKRWVWFDASNPDPNQLPDESHPVHNPSPMSIVS-----SYDANILVYI 422

```

```

OY 203 GSTAVGPASEWQKPT---GGTLGNTFR-----AVRDTGEIWRKQTLPRDWMDEC 254
Db 423 PMGV-----GTPDWMGDRKTSERAPGVALNMDTKLAFYOTVHHNDMDL 473
OY 255 TPEMAVTVNDVOPSTEMEGIOSINPNNAATGERVLYGPKCTTMMQCPAENG-EFLMAR 313
Db 474 PQGSLVNDVOTKDDGLVPAIYA-----PKTGDIVLDKRRKKEIVAP 517
OY 314 DT-----NYQMIIESIDENGIVTVNEDALKELDV-----EYD 346
Db 518 EEPVPOGAAPGHTPTQPMG---LTLRPKNPLNDSIDMGFTFDQECISYFHTLRYE 574
OY 347 -----VCPFLGGRDWPSPALNPDGSIYFTPLNNVCYDMAVYO----- 385
Db 575 GPPTPLSINGSLIFPDGLQMFEMGGIAVDPOQVAFANPISLPVSQLVYRGGNPLMPE 634
OY 386 -----EFTSMDEVNTSNVTKLPKGDML-----GRIDAIDISTGRTL 422
Db 635 ENAKGTGEGTLOHNYGTYAVNLHPFLDPVLLPFGIKMPCRTPPWGVAGIDIKTKKVY 694
OY 423 WSVERRAANYS-----PV-----LSTGGVYLF-NGDTRFERRALSOETG 460
Db 695 WQHNGTGRDSMNGSSLPPLPPIKIGVPSLGGPLSTAGNLGFLTASMDYYIRAYNLTTG 754
OY 461 ETLWQTRLATVAVSQAISYEVDGMOYVAIAGGVSYSYSGILNS 502
Db 755 KVLWODRLPAGAOATPTTVAINGKOYI-----VTYAGGHS 790

```

RESULT 15
B98314

Probable quinate dehydrogenase [imported] - Agrobacterium tumefaciens (strain C58, Cerec)
C:Species: Agrobacterium tumefaciens
C:Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 11-Jan-2002
C:Accession: B98314
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman, R.; Li, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.
A: Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tumefaciens
A:Reference number: A97359; PMID:11743194
A:Accession: B98314
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-809 <KUR>
A:Cross-references: GB:AE007870; PIDN:AAK90036.1; PID:g15160013; GSPDB:GN00170
C:Genetics:
A:Gene: AGR_L_2936
A:Map position: linear chromosome
C:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)

Query Match 10.1%; Score 271; DB 2; Length 809;
Best local similarity 21.6%; Pred. No. 1.6e-11;
Matches 152; Conservative 70; Mismatches 224; Indels 258; Gaps 30;

```

OY 8 WASGATLALAAPAF-----AQTPTVDLLANPPAGMISYQONQENY 51
Db 128 FAVAAVALAAGAGFAGFVAHPTVAFNGEPALTPVKPE-----TEQKNWEHYGNTSGGS 183
OY 52 RHPSPLOTITENVOQLDLYW-----ARGMQPKGVOTPLINDGVYLANPDDVTOA 102
Db 184 RFVALDEITRDNINKLEAMTYHTGDTPISPGANGAEDDOTPLQVGDVFLCTPHNNVYA 243
OY 103 IDAKTGDLIWE-----HR 115
Db 244 LDADGTETKWKAKINSKSSVMRCGLAYPDAPAKAPLKOPTAPGSTPYTAAYVAEGALCOR 303
OY 116 ROLPNTATLNIYA-----NG-----VTY 133
Db 304 RILMTNINELIALDADGACPDGFTNGRDALKIGMDADDPQYVLTSAFTLAGTIVVY 363
OY 134 AG--STCGYSPGCGFVSGHDSATGEELMRYFTIPRAGEBDETW-----GNDYEARMM 184

```

```

Db 364 GGRIDNNVQVMPGVMRGFDVYTGELRW-----AFDEGNPDITKLPKPGQTY-TRST 415
OY 185 TGAAGQITVDPTNLVH--YGSTAVGPASEFTQRTPGGTLGNTFRFAVRPDDGEIWRH 242
Db 416 PNWMASSYDPDLNTVFMFVPGSPSVDLYGATR--TPLDHKYGA-SMLADATGREGKMY 472
OY 243 QTLPRNNQOECEFFEMAVTVNDVOPSTEMEGIOSINPNNAATGERVLYGPKCTTMMQCPAENG-EFLMAR 313
Db 473 QTVHNDLMD-----FDVPMQPS-----FVDFPKADG-----TSVPALVFGTKAG 511
OY 298 TMMQFPAETG-----EFLMARDTNYQ-----MIESIDENGIVTVNEDALKELDV-----EYD 346
Db 512 QLYVLDRTAQOPLTKYEEVSVKANIIPNEPYAPTPQPRSVGMQDGAOTLFTPADMMGATPF 571
OY 335 DALIKEL--DVEYD-----VCPFLGGRDWPSPALNPDGSIYFTPLNNVCYDMAVYO----- 385
Db 572 DQILCRIAFKGRREGIYTAFTGDSLAFPGSLGGMMNGSLSTDPTTITVTVNDKRLGLW 631
OY 372 -----PL-----NNVCYDMAVDOEFTSMDEVNTSNVTKLPKGDML-----GRIDAIDISTGRTL 422
Db 632 IEMKAAPTAKAVASGESVNTGMGVPMKGTPTAAYVKNRFLSALGIPCQAP-----YGTM 687
OY 412 DAIDISTGRTLMSVERAANYS-----PV-----LSTGGVYLF-NGDTRFERRALSOETG 460
Db 688 TALDKTKQIKMOPVYGVTEDTGPIGIMGLPIPIGMPFLGTLTLAGGGLVFTLAGTQDY 747
OY 452 FRALSOETGETLWOTRLATVAVSQAISYE--VDGMOYVAIAGG 493
Db 748 LRAFDTATGKEVWKARLPVSGSGGPMSTKSPKTKQKVIVVISAAG 791

```

Search completed: May 24, 2002, 10:21:06
Job time: 344 sec

Fri May 24 11:27:26 2002

wallick-934-125.pep.rpr

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2002, 10:30:50 ; Search time 28.06 Seconds
(without alignments)
692.702 Million cell updates/sec

Title: WALICK-934-125.PEP
Sequence: 1 MKPTSLTMSAGALALLAAP.....GMOYVAINGGVSYGSLINS 502

Scoring table: BLAST62
Gapop 10.0, Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	479	17.8	739	1 DHET_ACEU	044002 acetobacter
2	475	17.7	742	1 DHET_ACEB	P18278 acetobacter
3	474.5	17.6	738	1 DHET_ACEP	P28016 acetobacter
4	461.5	17.2	757	1 DHET_GLUO	O05342 glucobacter
5	435.5	16.2	626	1 DHM1_METX	P16027 methylolact
6	435.5	16.2	626	1 DHM1_METX	P15279 methylolact
7	426.5	15.9	600	1 XOXF_PARD	P29968 paracoccus
8	424	15.8	631	1 DHM1_PARD	P12293 paracoccus
9	413	15.4	623	1 EXAA_PSEAE	O92417 pseudomonas
10	388.5	14.4	571	1 DHM1_METME	P38539 methylolact
11	339.5	12.6	766	1 DHG1_ECOLI	P15877 escherichia
12	315	11.7	801	1 DHGA_ACIKA	P03465 acetobacter
13	287.5	10.7	808	1 DHG1_ECOLI	P27175 glucobacter
14	250.5	9.3	809	1 DHG1_ECOLI	O59086 acetobacter
15	240	8.9	790	1 DHG1_ECOLI	O59086 acetobacter
16	134.5	5.0	827	1 YFGL_ECOLI	O60282 methanococcus
17	129.5	4.8	392	1 YFGL_ECOLI	P77774 escherichia
18	123.5	4.7	593	1 SPG2_STRSP	P19909 streptomyces
19	118.5	4.4	1355	1 AFSK_STRCO	P22705 anabaena sp
20	116.5	4.3	1355	1 AFSK_STRCO	P22705 anabaena sp
21	115.5	4.3	353	1 YXAL_BACSU	O42111 bacillus su
22	115.5	4.3	353	1 YXAL_BACSU	O42111 bacillus su
23	115.5	4.3	353	1 YXAL_BACSU	O42111 bacillus su
24	114	4.2	796	1 COPP_SCHPO	O42937 schizosacch
25	112	4.2	415	1 Y232_RICPR	O92482 rickettsia
26	112	4.1	3317	1 OMPB_RICPR	P96989 rickettsia
27	110.5	4.0	954	1 FLEY_CAUCR	P53365 rattus norv
28	108.5	4.0	3354	1 CADN_HUMAN	P15345 caulobacter
29	108.5	4.0	1323	1 RPOD_HUMAN	O91251 homo sapien
30	107.5	3.9	621	1 ASPA_AERSA	P28897 straberry
31	106	3.9	649	1 ACES_DRONE	P31339 aeromonas s
32	104	3.9	697	1 NAME_STRPN	P07140 drosophila
33					O54727 streptococc

ALIGNMENTS

RESULT ID	1	STANDARD	PRT	739 AA	
1	DHET_ACEU				
AC	044002; 007952;				014301 schizosacch
DT	15-DEC-1998 (Rel. 37, Created)				P80544 staphylococ
DT	15-DEC-1998 (Rel. 37, Last sequence update)				P46881 acetobacte
DT	15-JUL-1999 (Rel. 38, Last annotation update)				P06867 sus scrofa
DE	Alcohol dehydrogenase [acceptor] precursor (EC 1.1.99.8).				P27276 avian infec
GN	ADH.				P29336 streptococ
OS	Acetobacter europaeus.				P06654 streptococ
OC	Bacteria; Proteobacteria; alpha subdivision; Acetobacteraceae;				O88766 rattus norv
OC	Glucobacter.				O70138 mus musculu
OX	NCBI_TaxID=33995;				P30270 streptomyce
RN	[1]				P97260 cricetus
RP	SEQUENCE FROM N.A.				P54742 streptomyce
RC	SPRAIN-DES11 / DSM 6160;				
RA	Thurner C.A.K.;				
RL	Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.				
CC	-1- SUBUNIT: HETEROETRAMER (DEHYDROGENASE, CYTOCHROME, AND TWO				
CC	-1- COFACTOR: PO4 AND HEME (BY SIMILARITY).				
CC	-1- REDUCED ACCEPTOR				
CC	-1- CATALYTIC ACTIVITY: A primary alcohol + acceptor -> an aldehyde +				
CC	SMALLER UNKNOWN SUBUNIT(S) THAT FORMS THE ALCOHOL DEHYDROGENASE				
CC	COMPLEX (BY SIMILARITY).				
CC	-1- SUBCELLULAR LOCATION: MEMBRANE-BOUND, FACING THE PERIPLASMIC				
CC	SPACE (POTENTIAL).				
CC	-1- SIMILARITY: BELONGS TO THE BACTERIAL PO4 DEHYDROGENASE FAMILY.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/				
CC	or send an email to license@sib.ch).				
CC	-----				
CC	EMBL: X82894; CAA58066.1; -				
CC	EMBL: Y09480; CAA70688.1; -				
CC	HSSP: Q92477; IFLG.				
CC	InterPro: IPR001479; Bac_P00.				
CC	InterPro: IPR002372; Bac_P00_repeat.				
CC	InterPro: IPR000345; CytC_heme_bind.				
CC	Pfam: pf01011; Bacterial_P00_1; 6.				
CC	PROSITE: PS00364; BACTERIAL_P00_2; 1.				
CC	PROSITE: PS00190; CYTOCHROME_C_1.				
CC	DR OXIDOREDUCTASE; PO4; Heme; Periplasmic; Membrane; Signal.				
CC	KW SIGNAL				
CC	FT CHAIN 1 35				
CC	FT BINDING 651 739				
CC	FT BINDING 651 651				
CC	FT BINDING 654 654				
CC	FT BINDING 655 655				
CC	FT METAL 739 AA; 80944 MW; E681BB237ACB91F4 CRC64;				
CC	FT SEQUENCE				

Query Match

17.8% Score 479; DB 1; Length 739;

Best Local Similarity 25.9%, Pred. No. 2.9e-28;
Matches 151; Conservative 77; Mismatches 236; Indels 120; Gaps 15;

```

OY 10 SAGALLALAAPAFQVPTVDE-----LLANPAGEMISYGOQENVRHSPLQIT 61
DB 17 TGTGICALALLISYATMASADGOGATEAIHADDHGMNMTYGRISQORSPIDQINR 76
OY 62 ENVGLOLVARGMOPKQVYV-TPILHGVMTLANGVYQAIADAKTGLIWEHRRLP- 119
DB 77 SNVGNKLMWYLDIDTRNGEGTPLYDGMVATTWMMKAVDAATGATLMSYDPRVG 136
OY 120 NIA-----TLN----- 125
DB 137 NIADKGGCDYVNGAATWNGKYEGFEDRLIALDAKTGLIWSVNTIPPAELGKORSY 196
OY 126 -----LVANGVYVASTCOYSPFGC--FVSGHDSATGEBELMRNFTRAGEEDETWGN 177
DB 197 TVDGAPRIAGRYTIGN--GSGFEGARGFVTAFADETKRVDMREFETAPNKPEDHTASD 254
OY 178 D-----YEARWMTGA-----WQITYPVTNLVHGSTAVGPASETQRTGRT 221
DB 255 SVLMNKAYOTWSPGTGAMTNGGQGGVWMSIYDPAVDLYLVGVNGSGPMNYKTRSEKGD 314
OY 222 LVGTNRFAVRPDTEIGIWRHQTLPDRNMDOCTEFEMWVTVDVQPTSEBGLQSTNPA 281
DB 315 NLFSLIYALKPTEBYWVHFOETPMQMDFTSVQIMTIDLPLNGET----- 362
OY 282 ATGERRVLTGVPCKTGTWMOFDEGFEIAROTNIONMIESID-ENGIVYVEDAIKE 340
DB 363 -----BHVIYVAP-KNGFEYIIDAKTGFISGKNVYVNASGIDPRTGPRIPNPAALVTL 417
OY 341 LDVEYDVPFELGGRDPSAALNPDGSIYFIPLNVCYDMMAVDQFT-SMUYNTS--- 396
DB 418 TGKEMVYGLPDLGHNRAAFSPKGTGLVYIPAOQVPLTYTNOVGFTPHDPSMGLDM 477
OY 397 NVTLKLPK-----KDMIGRIDAIDISTGRTLSVEEAANYSPLYSTGGVLENGGT 448
DB 478 NKVGIPSPKQAFVYVDLKGKIVAMDPOKQAMRNVDHKGWNGGILATGDLFOGLA 537
OY 449 DRYFRALSOETGLMOTRLATVYASQALSYVDMGYVAAAG 492
DB 538 NGEFHAYDATNGSDLFHFAADSGIAPVYTLANGKQYVAVEVG 581

RESULT 2
DHET_ACEAC STANDARD; PRT; 742 AA.
AC F18278;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Alcohol dehydrogenase [acceptor] precursor (Ec 1.1.99.8).
GN ADHA OR ADH1.
OS Acetobacter aceti.
OC Bacteria; Proteobacteria; alpha subdivision; Acetobacteraceae;
OC Acetobacter.
OX NCBI_TaxID=435;
RN [1]
RP SEQUENCE FROM N. A., AND SEQUENCE OF 36-50.
RX MEDLINE=89255070; PubMed=2722742;
RA Inoue T., Sunagawa M., Mori A., Imai C., Fukuda M., Takagi M.,
RA Yano K.,
RT "Cloning and sequencing of the gene encoding the 72-kilodalton
RT dehydrogenase subunit of alcohol dehydrogenase from Acetobacter
RT aceti."
RL J. Bacteriol. 171:3115-3122(1989).
RN [2]
RP 3D-STRUCTURE MODELING.
RX MEDLINE=9528964; PubMed=7772016;
RA Cozier G.E., Giles I.G., Anthony C.;
RT "The structure of the quinoprotein alcohol dehydrogenase of
RT Acetobacter aceti modelled on that of methanol dehydrogenase from
RT Methylobacterium extorquens."

```

```

RL Biochem. J. 308:375-379(1995).
CC -1- CATALYTIC ACTIVITY: A primary alcohol + acceptor = an aldehyde +
CC reduced acceptor.
CC -1- COFACTOR: POQ AND HEME.
CC -1- SUBUNIT: Tetramer of non identical chains (DEHYDROGENASE,
CC CYTOCHROME, AND TWO SMALLER UNKNOWN SUBUNITS).
CC -1- SUBCELLULAR LOCATION: Periplasmic.
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POQ DEHYDROGENASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; D90004; BA14058.1; -
DR PIR; J50326; J50326.
DR HSSP; 0924J7; 1FIC.
DR Interpro; IPR001479; Bac_POQ.
DR Interpro; IPR002372; Bac_POQ_repeat.
DR Pfam; Pfold011; Bacterial_POQ.
DR PROSITE; PS00363; BACTERIAL_POQ_1; 1.
DR PROSITE; PS00364; BACTERIAL_POQ_2; 1.
DR PROSITE; PS00190; CYTOCHROME_C_1.
KW Oxidoreductase; POQ; Heme; Periplasmic; Membrane; Signal.
FT CHAIN 1 35
FT ACT_SITE 36 742 ALCOHOL DEHYDROGENASE [ACCEPTOR].
FT BINDING 649 649 HEME (POTENTIAL).
FT BINDING 652 652 HEME (COVALENT) (BY SIMILARITY).
FT METAL 653 653 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
SQ SEQUENCE 742 AA; 81521 MW; 9C6C9268DABB825A CRC64.

```

Query Match 17.7%; Score 475; DB 1; Length 742;
Best Local Similarity 26.0%; Pred. No. 5.9e-28;
Matches 160; Conservative 87; Mismatches 202; Indels 166; Gaps 23;

```

OY 11 AGALLALAAPAFQVPTVDELT--ANPAGEMISYGOQENVRHSPLQITTNVGLQ 68
DB 23 AALPVAAPVADGOGNGTEAIHADDHGMNMTYGRISQORSPIDQINRSMVGLK 82
OY 69 LVNARGMOPKQV-TPILHGVMTLANGVYQAIADAKTGLIWEHRRLP-NIA----- 122
DB 83 LGGVYTLNTRNGEGTPLYDGMVATTWMMKAVDAATGATLMSYDPRVGNDKGC 142
OY 123 --TLN----- 126
DB 143 CDTVNRGAGYRNGKVFMGTFDGRVAAADAKTKRVAVNTIPADASLGORSYVDGAVR 202
OY 127 VANGVYVASTCOYSPFGC--FVSGHDSATGEBELMRNFTRAGEED----- 172
DB 203 VAKGLVLTGN--GGAERFARGFVSAFPAETGKLWRYTVPNNKNEPDAASDNLANKA 260
OY 173 -ETWANDYEAQMT-----GAMQITYPVTNLVHGSTAVGPASETQRTGRTGGLY--- 223
DB 261 YKTMGP--KGAMVROGGGTVWDSLVYDPSDLY--LAVG-----NGSPMNYKRSR 309
OY 224 --GTN-----TFEFAVRPDTEIGIWRHQTLPDRNMDOCTEFEMWVTVDVQPTSEBGLQSI 277
DB 310 GIGSNLFLGSIYALKPTEBYWVHFOETPMQMDFTSVQIMTIDLMPYK----- 358
OY 278 NPNMAGE-RVLTGVPCKTGTWMOFDEGFEIAROTNIONMIESID-----EN 327
DB 359 -----GEMHVIYVAP-KNGFEYIIDAKTGFISGKNVYVNASGIDPRTGPRIPNPAALVTL 411
OY 328 GIYTVVEDAILKELDEYDVPFELGGRDPSAALNPDGSIYFIPLNVCY----- 378
DB 412 GLYTLNG-----KFWGVI-PGLGANNFMAMVSPKTHLVYIPAHQIPFGYKNQVGGF 463

```

OY 379 -----DMAVDEFTSMQVYNTSKLPCKDMIGRIDISTGRFLMSVERAANTSP 434
 DB 464 KPHADSMVNGIDMTKNGLPDPE-ARTAYIKADLHGLIANDPVPKMETWIKIDHKGFWNG 522
 OY 435 VLTSGGVYFNGGNDYRFRALSOETGETLMOTRLATVASSOALSYEDQMYYA----- 488
 DB 523 ILATGGDLLEFOLANGEFHAYDATNGSDLYKFDAGSIIAPMTYSNGQYAVGVNG 582
 OY 489 ----TAGGVSYSG 499
 DB 583 GYPISMGVGRISG 597

RESULT 3
 DHET_ACEPO STANDARD: PRT: 738 AA.
 AC P28036;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DT Alcohol dehydrogenase [acceptor] precursor (EC 1.1.99.8).
 GN ADHA.
 OS Acetobacter polyoxogenes.
 CC Bacteria; Proteobacteria; alpha subdivision; Acetobacteraceae;
 CC Acetobacter.
 CC NCBL_TaxID=439;
 RN [1]
 RP SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.
 RC STRAIN-NB11028;
 RX MEDLINE=91159482; PubMed=2001402;
 RA Tamaki T., Fukaya M., Takemura H., Tayama K., Okumura H.,
 RA Kawamura Y., Nishiyama M., Horinouchi S., Beppu T.;
 RT "Cloning and sequencing of the gene cluster encoding two subunits of
 RT membrane-bound alcohol dehydrogenase from Acetobacter polyoxogenes."
 RT Biochim. Biophys. Acta 1088:292-300(1991).
 CC -1- FUNCTION: DEHYDRATION OF PRIMARY ALCOHOLS (EXCEPTION: METHANOL).
 CC -1- CATALYTIC ACTIVITY: A primary alcohol + acceptor = an aldehyde +
 CC reduced acceptor.
 CC -1- COFACTOR: POO AND HEME.
 CC -1- SUBUNIT: HETEROTETRAMER (DEHYDROGENASE, CYTOCHROME, AND TWO
 CC SMALLER UNKNOWN SUBUNITS) THAT FORMS THE ALCOHOL DEHYDROGENASE
 CC COMPLEX.
 CC -1- SUBCELLULAR LOCATION: MEMBRANE-BOUND, FACING THE PERIPLASMIC
 CC SPACE (POTENTIAL).
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POO DEHYDROGENASE FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: D00635; BAA00528.1;
 CC DR PIR: S14270; S14270.
 DR HSP: Q924J7; 1F7G.
 DR InterPro: IPR001479; Bac_POO.
 DR InterPro: IPR002372; Bac_POO_repeat.
 DR InterPro: IPR000345; CYC_Heme_bind.
 DR Pfam: PF01011; Bacterial_POO; 6.
 DR PROSITE: PS00363; BACTERIAL_POO_1; 1.
 DR PROSITE: PS00364; BACTERIAL_POO_2; 1.
 DR PROSITE: PS00190; CYTOCHROME_C; 1.
 DR PROSITE: POO; Heme; Periplasmic; Membrane; Signal.
 DR OXidoreductase; 1.
 DR SIGNAL 35
 DR CHAIN 36 738
 DR BINDING 650 630
 DR BINDING 653 653
 DR METAL 654 654
 DR SEQUENCE 738 AA; 80840 MW; 1E2B6ED7BCD92AF6 CR64;
 SQ

Query Match 17.6%; Score 474.5; DB 1; Length 738;
 Best Local Similarity 25.6%; Pred. No. 6,4e-28;
 Matches 151; Conservative 77; Mismatches 229; Indels 133; Gaps 15;
 OY 10 SAGALALLAAPFAQYTPVTDE-----LIANPPGEWISYSGONENYRHSPLQIT 61
 DB 17 TAGTICALISGATWASADDOGATGEALIHADHPGNMNYGRTYSDQYSPLDQINR 76
 OY 62 ENVGQOLYVARBMQNGKQYV--PRLHDGVMYLANPGVQAIQAKTGDILWEHRRLP 119
 DB 77 SNGNKLAMVLLDITNRQGEPTPLVDGVMYATINMSKAVDAATGKLMSYDPKYPG 136
 OY 120 NIA-----TLN----- 125
 DB 137 NIADGGCDIVNRGAAYNGKVFYFEDGRLIALDAKTKLWVSNTIPEAELEKORSY 196
 OY 126 -----YANGVIYAGSTCOYSPGC--FVSGHDSATGEELRNFIPLAGEED----- 172
 DB 197 IVDGAPRIAKGVILIGN--GSGEFGARFVSARFADETGKVDWRFETVPPKNEPDAIDS 254
 OY 173 -----ETWGNDEYARMWTCG-----WGQIYDPVTMLVHGSYAVGPASETOR 215
 DB 255 VLMNKAYQFWS-----PTGAMTRQGGGGTWDSDIYDDVADLVYLGNGSPMYKRYR 307
 OY 216 GTPGGLTYGNTNRFAPVPTGEIWRHQTLPDRWDDECFEMKATVNDYQSTEMEGIQ 275
 DB 308 SEKGDNLFGLSIVALKPRTGEYVWHFOETPMQDMDFTSDQINTLDP----- 357
 OY 276 SINPNATGERRVLTGVPCKTGMQFDEGEFLMARDTYQNMIESID-ENGIVTNE 334
 DB 358 -----NGETRHIVHARKNGEFTYIDAKTGEFISGNYYVWMAASGLDKTRPIYNP 410
 OY 335 DALLKELDYVDCPTFLGGRDPSALNPDGSIYFIPLANVCYDMAVDEFT-SMDY 393
 DB 411 DALYTLTGKWEYGLPGDLGSHNFAAMAFSPKGLVYIPAOVPELYTQVGFPHRPSW 470
 OY 394 NTS---NYTKLPSC-----KDMIGRIDADISTGRFLMSVERAANYSPVLTSGGV 442
 DB 471 NLGDMNKVGIIPDSPEAKQAFVDLKWIVAMPQQAEMWVDHDKPMNGGIIATGGDL 530
 OY 443 LFNGTIDYFRALSOETGETLMOTRLATVASSOALSYEDQMYYA IAG 492
 DB 531 LFOGLANGEFHAYDATNGSDLYKFDAGSIIAPMTYSNGQYAVGVNG 580
 RESULT 4
 DHET_GLUOX STANDARD: PRT: 757 AA.
 AC 003542;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Alcohol dehydrogenase [acceptor] precursor (EC 1.1.99.8) (G3-ADH
 DE subunit 1).
 GN ADHA.
 OS Gluconobacter oxydans (Gluconobacter suboxydans).
 CC Bacteria; Proteobacteria; alpha subdivision; Acetobacteraceae;
 CC Gluconobacter.
 CC NCBL_TaxID=442;
 RN [1]
 RP SEQUENCE FROM N.A. AND SEQUENCE OF 36-50.
 RC STRAIN-IFO 12528;
 RX MEDLINE=97208225; PubMed=9055427;
 RA Kondo K., Horinouchi S.;
 RT "Characterization of the genes encoding the three-component membrane-
 RT bound alcohol dehydrogenase from Gluconobacter suboxydans and their
 RT expression in Acetobacter pasteurianus."
 RT Appl. Environ. Microbiol. 63:1131-1138(1997).
 RL Appl. Environ. Microbiol. 63:1131-1138(1997).
 CC -1- FUNCTION: DEHYDRATION OF PRIMARY ALCOHOLS (EXCEPTION: METHANOL).
 CC -1- CATALYTIC ACTIVITY: A primary alcohol + acceptor = an aldehyde +
 CC reduced acceptor.
 CC -1- COFACTOR: POO AND HEME.

CC -1- SUBUNIT: HETEROTRIMER (DEHYDROGENASE, CYTOCHROME AND PROTEIN
 CC ADHS), THAT FORMS THE ALCOHOL DEHYDROGENASE COMPLEX.
 CC -1- SUBCELLULAR LOCATION: MEMBRANE-BOUND, FACING THE PERIPLASMIC
 CC SPACE (POTENTIAL).
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POQ DEHYDROGENASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; D86375; BAA19753.1; -
 DR HSSP; Q924J7; LF1G;
 DR InterPro: IPR001479; Bac_PoQ.
 DR InterPro: IPR002372; Bac_PoQ_repeat.
 DR Pfam: PF01011; Bacterial_PoQ; 6.
 DR PROSITE: PS00363; BACTERIAL_POQ_1; 1.
 DR PROSITE: PS00364; BACTERIAL_POQ_2; 1.
 DR PROSITE: PS00190; CYTOCHROME_C; 1.
 DR OXidoreductase; POQ; Heme; Peptide; Signal.
 KW SIGNAL 1 34
 FT CHAIN 1 34
 FT MOD_RES 35 757 ALCOHOL DEHYDROGENASE [ACCEPTOR].
 FT ACT_SITE 342 342 BASE (POTENTIAL).
 FT BINDING 653 653 PYRROLIDONE CARBOXYLIC ACID.
 FT BINDING 656 656 HEME (COVALENT) (BY SIMILARITY).
 FT METAL 657 657 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
 SQ SEQUENCE 757 AA; 82968 MW; 3989F90E3B947581 CRC64.

Query Match 17.2%; Score 461.5; DB 1; Length 757;
 Best Local Similarity 25.5%; Pred. No. 6.2e-27;
 Matches 156; Conservative 87; Mismatches 225; Indels 143; Gaps 21;

QY 7 LMSAGALALIAA--PAFAVY--TPVTELLANPAGEMISYGOENYRHSPLTOIT 61
 DB 16 LLSAALALSAVPAFAFQEDTGATITSSDNGHP--DMLSTYGRSYEQSPSLDINT 74
 QY 62 ENTGOLIVWARGMOPGVYV--TPIIHGVYVLANPVDYIQAIDAKGDLWEHRQLP 119
 DB 75 ENVGKILKLMHVDLTNRGOESTPLIVGVVATTNMSKMAALDAVGLKLMYSRVP 134
 QY 120 NTA-----TLNIVA--NGVIVAGS----- 136
 DB 135 NIADRCCTVSRGAAYNGKAVYFETEDGLIALDAKGLVMSVYTIPEAOLGHORSY 194
 QY 137 TGOYSP-----FGC--FVSGHDSATGEELMRNYETIRAGEBD----- 172
 DB 195 TYDGAAPRAKRVLLINGNGAEFGARFVSAPDAETSKLDRFTVNPENKKGDAASDDI 254
 QY 173 -----ETWGNDEYBARWMTG--AMGQITDPTVNLVHGSGAVGASFTQGTGGTLY 223
 DB 255 LMSKAYPTMGKNGAMKQGGGGTWSLVYDPTDLYLVGCVGSPWNNKFFSEBKGNL 314
 QY 224 GTNRFAVRPDTEIYVRHQTLPDNDQDCPEMAYTVNDVQSTEWEGLOSTINPNAT 283
 DB 315 FLGSIIVAINPDTRKYVHFQETPMEDMTVSQOILMTLMPV-----N 357
 QY 284 GE--RRVLTGVPCSTGTMQPDATGTEFLMARDTNTQNNIESIDE--NGLVTVNEDAILKEL 341
 DB 358 GEMRHVIVAP--KNGFFYITIDAKTGKFTYIGKPYEYENANGLDPTVGTGPNVVDALMTLT 416
 QY 342 DVEYDVCPTFLGDRWPSAALNPDGSIYFIPNNV-----CYDMAVDOEFT 388
 DB 417 GKFWLIGPELIGHNFAAAVSPKTKLYITPAQOPLIYDGQKGFAYDANNLGIDNN 476
 QY 389 SMDVYTSVWTLPKPKMIGIDAIDISTGRTL--WSVERAAANYS-----PVLSF 438
 DB 477 KIGLPDNDPEHVAKKDF-----LVYLGKGVAVAMPPEKMAFATPINHGPNWGLLAT 530

QY 439 GGVLENGGTDRYFRALSOETGETLMQRLNVAAGSISYVDGMOYAI----- 489
 DB 531 AGNVIFQGLANGFEHAYATATNNDLXSPQASAIAPPYTANGKQYVAEVMGCIIP 590
 QY 490 -AGGGSYSYG 499
 DB 591 FLXGVYARISG 601

RESULT 5
 DHML_METEX STANDARD; PRT; 626 AA.
 AC P16027;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-APR-1990 (Rel. 14, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Methanol dehydrogenase subunit 1 precursor (EC 1.1.99.8) (MDH large
 DE alpha subunit) (MDH).
 GN MOXF.
 OS Methylobacterium extorquens.
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Methylobacterium group; Methylobacterium.
 OX NCBI_Taxid=408;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=AM1 / NCIB 9133;
 RX MEDLINE=90337342; PubMed=2116368;
 RA Anderson D.J., Morris C.J., Nunn D.N., Anthony C., Lidstrom M.E.;
 RT "Nucleotide sequence of the Methylobacterium extorquens AM1 *mox* and
 RL gene 90:173-176(1990)."
 RN [2]
 RP SEQUENCE OF 28-53.
 RC STRAIN=AM1 / NCIB 9133;
 RX MEDLINE=89350892; PubMed=2504152;
 RA Nunn D.N., Day D., Anthony C.;
 RT "The second subunit of methanol dehydrogenase of Methylobacterium
 RL extorquens AM1.";
 RN Biochem. J. 260:857-862(1989).
 RN [3]
 RP DISULFIDE BONDS.
 RX MEDLINE=95384759; PubMed=7656012;
 RA Blake C.E.F., Ghosh M., Harlos K., Avezoux A., Anthony C.;
 RT "The active site of methanol dehydrogenase contains a disulphide
 RL bridge between adjacent cysteine residues.";
 RN Nat. Struct. Biol. 1:102-105(1994).
 RN [4]
 RP X-RAY CRYSTALLOGRAPHY (1.94 ANGSTROMS).
 RX MEDLINE=95253818; PubMed=7735834;
 RA Ghosh M., Anthony C., Harlos K., Goodwin M.G., Blake C.;
 RT "The refined structure of the quinoprotein methanol dehydrogenase
 RL from Methylobacterium extorquens at 1.94 A.";
 RN Structure 3:177-187(1995).
 CC -1- CATALYTIC ACTIVITY: A primary alcohol + acceptor = an aldehyde +
 CC reduced acceptor.
 CC -1- COFACTOR: POQ.
 CC -1- SUBUNIT: THE HOMOZYME MDH HAS AN ALPHA-2/BETA-2 CONFIGURATION
 CC OF A LARGE ALPHA AND A SMALL BETA SUBUNIT.
 CC -1- SUBCELLULAR LOCATION: PERIPLASMIC, BUT ASSOCIATED WITH INNER
 CC MEMBRANE.
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POQ DEHYDROGENASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M31108; AAA25380.1; -
 DR PIR; S07908; S07908.

PIR: J00706; J00706.
 HSP: P38539; 4AAH.
 DR InterPro: IPR001479; Bac-PQQ-repeat.
 DR InterPro: IPR002372; Bac-PQQ-repeat.
 DR Pfam: PF01011; Bacterial_PQQ_7.
 DR PROSITE: PS00363; BACTERIAL_PQQ_1; 1.
 DR PROSITE: PS00364; BACTERIAL_PQQ_2; 1.
 DR Oxidoreductase; PQQ; Signal; Methanol utilization; Periplasmic.
 KW SIGNAL 1 27
 FT CHAIN 28 626 METHANOL DEHYDROGENASE SUBUNIT 1.
 FT DISULFID 130 131
 FT DISULFID 413 442
 FT ACT_SITE 330 330
 FT SEQUENCE 626 AA; 68434 MW; 649880DAFD2AD34C CRC64;
 SO
 Query Match 16.2%; Score 435.5; DB 1; Length 626;
 Best local similarity 24.6%; Pred. No. 4.3e-25;
 Matches 148; Conservative 91; Mismatches 208; Indels 155; Gaps 23;
 Oy 10 SAGALALLA-AP-AFAOVPPTDELINAPAGE-WISYGOENYRHSPLTQITTEVGO 66
 Db 7 SVSALMALALAPALSSAVANDKLEISKDDNMWPKNSNNSELDKQIKKKNVQ 66
 Oy 67 LQIYW--ARGMOPGVQVYPLIHGVMY-----LANPGDY----- 99
 Db 67 LRPAMTFSTGLNGH-EGAPLVYDGKMYHTSPNNTFALGLDDPGLIMODKPKONPAA 125
 Oy 100 -----IOAIDAKTGDLIWEHRQLPNIATLN 125
 Db 126 RAVACCDLVNGLAYWPGDGKTPALIKTQLDGNVALNAETGEVWK----- 173
 Oy 126 IVANGVIAGSTCOYSPF-----GC-----FVSGHDSATGEELMRYFI----- 164
 Db 174 -VENSIDIKVSTLTIAPVYVKDVIIGSSGAEIGVRYLAVYKGEQVRAVATGPKD 232
 Oy 165 -----PRAGEE--GDETWNDEYARWMTGA--WQOITYPVNTLVHGSTAV 207
 Db 233 DLLLADDFNKNNAHYGOKGLGTATWED--AWKIGGTWGMVAYDPGNTLITGTGNP 289
 Oy 208 GPASSETGCTPGGLYGTNTRFAVRPDTGEIWRHQLPRDNMDOCTEEMVTVNDVOP 267
 Db 290 APNNEETMR--PGDNKW--TWITIGRADDTGEAKFGYOKTPHEMPC-----YAGVNVMM 338
 Oy 268 STEMEGLSINPNAATGERRVLGVPCKTGTVMQDPAETGEELNAR--DTNYONMIESI 324
 Db 339 LSEKAD-----KDGKAKRLTLHPDRNGIYTLTDTGALVSANKLDDT--VNVFESV 388
 Oy 325 DENGIYVNEDELALKELD--VEYDVCTPLFGGRDWPSSALNPDGSIYFIPLNNVCYDMAV 383
 Db 389 DIKTGOVPRDEYGTTRMDLAKDICPSAMGYHNOCHSDYPRKELEFPGINHCIDWEPF 448
 Oy 384 DOEFMSDVNTSNTKLPPEK-----DMIGRIDAIDISTGRITLMSVERAANYSPLS 437
 Db 449 MLPRYAGGFVFGATILNMPGPKGRONYEGLQIKALYNALITGDYKKEKEREAVNGTGA 508
 Oy 438 TGGVULFNGGDRRFRALSOETGETLWOTRLATVASSQALSYEVNDQYVALAAGGVSYG 497
 Db 509 TAGDILVFETGLDGLKANDSDTGLMKFKIPSGALIGTGMPTVYTHKQYVAL-----YTG 563
 Oy 498 SG 499
 Db 564 VG 565
 RESULT 6
 DHM1_METOR STANDARD: PRT: 626 AA.
 AC P15279;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-APR-1990 (Rel. 14, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Methanol dehydrogenase subunit 1 precursor (EC 1.1.99.8) (MDH large

DE alpha subunit) (MDH).
 GN MOXE.
 OS Methylobacterium organophilum XX.
 OS Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Methylobacterium group; Methylobacterium.
 OC NCBI_taxid=410;
 OX (1)
 RN SEQUENCE FROM N.A., AND SEQUENCE OF 29-43.
 RP STRAIN=ATCC 27866 / DSM 246 / NCIB 11278;
 RX MEDLINE=89008094; Pubmed=2459109;
 RA Machlin S.M., Hanson R.S.;
 RT "Nucleotide sequence and transcriptional start site of the
 RT Methylobacterium organophilum XX methanol dehydrogenase structural
 RT gene";
 RL J. Bacteriol. 170:4739-4747(1988).
 RL J. CATALYTIC ACTIVITY: A primary alcohol + acceptor = an aldehyde +
 CC reduced acceptor.
 CC -1- COFACTOR: PQQ.
 CC -1- SUBUNIT: THE HOLOENZYME MDH HAS AN ALPHA-2/BETA-2 CONFIGURATION
 CC OF A LARGE ALPHA AND A SMALL BETA SUBUNIT.
 CC -1- SUBCELLULAR LOCATION: PERIPLASMIC, BUT ASSOCIATED WITH INNER
 CC MEMBRANE.
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL PQQ DEHYDROGENASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation on
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: M2629; AAA50289.1;
 DR HSP: P38539; 4AAH.
 DR InterPro: IPR001479; Bac-PQQ-repeat.
 DR InterPro: IPR002372; Bac-PQQ-repeat.
 DR Pfam: PF01011; Bacterial_PQQ_7.
 DR PROSITE: PS00363; BACTERIAL_PQQ_1; 1.
 DR PROSITE: PS00364; BACTERIAL_PQQ_2; 1.
 DR Oxidoreductase; PQQ; Signal; Methanol utilization; Periplasmic.
 KW SIGNAL 1 28
 FT CHAIN 29 626 METHANOL DEHYDROGENASE SUBUNIT 1.
 FT DISULFID 130 131
 FT DISULFID 413 442 BY SIMILARITY.
 FT ACT_SITE 330 330 BASE (POTENTIAL).
 FT SEQUENCE 626 AA; 68677 MW; 8768F6B8371E5DF CRC64;
 SO
 Query Match 16.2%; Score 435.5; DB 1; Length 626;
 Best local similarity 24.6%; Pred. No. 4.3e-25;
 Matches 148; Conservative 91; Mismatches 208; Indels 155; Gaps 23;
 Oy 10 SAGALALLA-AP-AFAOVPPTDELINAPAGE-WISYGOENYRHSPLTQITTEVGO 66
 Db 7 SVSALMALALAPALSSAVANDKLEISKDDNMWPKNSNNSELDKQIKKKNVQ 66
 Oy 67 LQIYW--ARGMOPGVQVYPLIHGVMY-----LANPGDY----- 99
 Db 67 LRPAMTFSTGLNGH-EGAPLVYDGKMYHTSPNNTFALGLDDPGLIMODKPKONPAA 125
 Oy 100 -----IOAIDAKTGDLIWEHRQLPNIATLN 125
 Db 126 RAVACCDLVNGLAYWPGDGKTPALIKTQLDGNVALNAETGEVWK----- 173
 Oy 126 IVANGVIAGSTCOYSPF-----GC-----FVSGHDSATGEELMRYFI----- 164
 Db 174 -VENSIDIKVSTLTIAPVYVKDVIIGSSGAEIGVRYLAVYKGEQVRAVATGPKD 232
 Oy 165 -----PRAGEE--GDETWNDEYARWMTGA--WQOITYPVNTLVHGSTAV 207
 Db 233 DLLLADDFNKNNAHYGOKGLGTATWED--AWKIGGTWGMVAYDPGNTLITGTGNP 289
 Oy 208 GPASSETGCTPGGLYGTNTRFAVRPDTGEIWRHQLPRDNMDOCTEEMVTVNDVOP 267

Db 290 AFNMTMR--PQDNKWM-TMTLEGRDADTGEAKRYOAKTPHEDWDYAGVNM-----P 339
 QY 268 STEMEGLQSIQINPNATGERRVLTGVPCKTGMQFPAEGEREFLMAR--DTYVQWMISSI 324
 Db 340 SPQKX-----KDGTRKLTHPRNGIVYTLDRTOALYSANKLDDT--VAVFKTV 388
 QY 325 DENGIVYVNEDELLKEID-VEDVCPFLGGRDWPSSAALNPDSCGIFLPLNNVCYDMAV 383
 Db 389 DLKGTGQVPRDEGCTRMHDIAKDVCPSSAMGYNHGSDIDPKRELFEFGINHICDMEPE 448
 QY 384 DQETSMQVNTSNTVTLPRGK-----DMGRIDALIDISTGRTLSVERAANYSPVLS 437
 Db 449 MLPFRAGOFVGAATLNNYRQPKKDRQNYGLQIKAYNAITGSKYEMEREFVWGTGLA 508
 QY 438 TGGGVLENGTDRFRLASQETGETLMTQRLATVASGOAISYEDVMQYVLAAGGVSYG 497
 Db 509 TAQDLVFTGLDYLKARDSDTGLIMKFKIPSGAIGVPMYTHKGYVAI-----YVG 563
 QY 498 SG 499
 Db 564 VG 565

RESULT 7
 XOXF_PARDE

ID XOXF_PARDE STANDARD: PRT; 600 AA.

AC P29968;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Putative dehydrogenase XOXF precursor (EC 1.1.99.-).
 GN XOXF.
 OS Paracoccus denitrificans.
 OC Bacteria; Proteobacteria; alpha subdivision; Rhodobacter group.
 OX NCBI_TaxID:266;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Harms N.;
 RL Submitted (Aug-1995) to the EMBL/GenBank/DBJ databases.

RP SEQUENCE OF 380-600 FROM N.A.
 RX STRAIN-PD 1235;
 RA MEDLINE-92041583; PubMed-1657873;
 RA Ras J., Reijnders W.N.M., Van Spanning R.J.M., Harms N., Oltmann L.F.,
 RT "Isolation, sequencing, and mutagenesis of the gene encoding
 RT cytochrome c551 of Paracoccus denitrificans and characterization of
 RL the mutant strain".
 RL J. Bacteriol. 173:6971-6979(1991).
 CC -1- COPACTOR: P00 (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL P00 DEHYDROGENASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb.sib.ch/announce/>
 CC or send an email to license@sib.ch).

DR EMBL: U34346; AAC44555.1;
 DR EMBL: M75583; AAA25574.1;
 DR PIR: A41378; A41378.
 DR HSSP: P38539; AAAH.
 DR InterPro: IPR002372; Bac_P00_repeat.
 DR Pfam: PF01011; Bacterial_P00; 7.
 KM Oxidoreductase; P00; Signal.
 FT CHAIN 1 21
 FT ACT_SITE 22 600 POTENTIAL.
 FT 318 318 PUTATIVE DEHYDROGENASE XOXF.
 FT SEQUENCE 600 AA; 65159 MW; DCA996F1BC5A3CE CRC64;

Query Match
 Best Local Similarity 25.8%; Score 426.5; DB 1; Length 600;
 Matches 156; Conservative 82; Mismatches 214; Indels 133; Gaps 26;

QY 10 SAGALALLAAPAFQVPTDELALNP-----PAGEATSTGQNDENTRHSPLOTITTE 62
 Db 6 NCACALIMSGTAA-----LANEQRGRORQAPQAWIDMGVANTRYSTLQINKD 56
 QY 63 NVGQLOLVW--ARGMGKQVQVPLIHGVYLANP-GDVQIAIDAK-TGDLMEHR-RQ 117
 Db 57 NVKDLRAWMTSTVLGH-ESSPLVIGVMYVHPFPRVFPALDNLNGKILMKRYEQQ 115
 QY 118 LPN-IATLNT-----VANGYVAGSTCOXSPGCFPSGSDATGELMRNF-IPRA 167
 Db 116 DPNVIAVMSQDYVYRGLSYADGMILLGQA-----DTIVVALDNTSGEYKSTRIGDGI 169
 QY 168 GE-----ECDETW-----GNDYEA----- 181
 Db 170 GETLTATVVPKQVLYVSGEYVGRGMALNLTDSSEAKMAMSTGPPDELLVDPETT 229
 QY 182 -----RWMTGA-----WQITVDPVNTLVHGVSTAVGASFGQGTG 219
 Db 230 THLGKPIGADSSLSNMEGDQMOIGGTLIGWFSYDPLNLVYIGTGNPSTWNSOR-BG 287
 QY 220 GYLXGTNRFVAVRPTGELVNRHOTLPDNDQCTEEMAVTN--VDVQSTEMEGLQSI 277
 Db 288 DNKW-SMTIMARDADTGAKKFFQMTPHDEWDYGVNEMILTNQVDCG----- 335
 QY 278 NPNATNGERYLVTPCKCTGTMOQFPAEGEREFLMARDTYVQWMISSI 332
 Db 336 -----ERKLTTFHD-RNGIAYTMDRETGELLVAE-KYDPVYNTTGVDDPNS-ETY 384
 QY 333 NEDALIKELDYED-----YCPFLGGRDWPSSAALNPDSCGIFLPLNNVCYDMAV 384
 Db 385 GRPAVVAEYSTAQGEDENTGCPALGKQDQPAEFPKTNLFVPTNCHMDYBERF 444
 QY 385 QETSMQV--NTSNTVTLPRGKMTGRIDAIDISTGRTLSVERAANYSPVLS 442
 Db 445 VAYTAGQPIYGATLSMPAPANSHGSGNFIAMHNTTGKISVPEPSVMSGALATAGDV 504
 QY 443 LFNQGTDRFRLASQETGETLMTQRLATVASGOAISYEDVMQYVLAAGGVSYG 497
 Db 505 VFYGTLEGLYKPYDAGTGELKFKFTPSGIIIGNVYTEHGKQYVGLISGVGAGIGLA 564
 QY 498 SG 502
 Db 565 AGILT 569

RESULT 8
 DHML_PARDE

ID DHML_PARDE STANDARD: PRT; 631 AA.

AC P12293;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Methanol dehydrogenase subunit 1 precursor (EC 1.1.99.8) (MDH large
 DE alpha subunit) (MDH).
 GN MOXF.
 OS Paracoccus denitrificans.
 OC Bacteria; Proteobacteria; alpha subdivision; Rhodobacter group.
 OX NCBI_TaxID:266;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 33-49.
 RA MEDLINE-87307969; PubMed-3114231.
 RA Harms N., de Vries G.E., Maurer K., Hoogenblik J., Stouthamer A.H.,
 RT "Isolation and nucleotide sequence of the methanol dehydrogenase
 RT structural gene from Paracoccus denitrificans".
 RL J. Bacteriol. 169:3969-3975(1987).
 CC -1- CATALYTIC ACTIVITY: A primary alcohol + acceptor -> an aldehyde +


```

CC reduced acceptor.
CC -1- COFACTOR: POQ.
CC -1- SUBCELLULAR LOCATION: Periplasmic.
CC -1- MISCELLANEOUS: MDH IS THE MAJOR PROTEIN IN THE CELL DURING GROWTH
CC ON METHANOL (IN P.DENITRIFICANS MDH CONSTITUTES UP TO 15% OF THE
CC TOTAL CELL PROTEIN).
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POQ DEHYDROGENASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: M17339; AAA8366.1; -.
CC HSSP: P38539; AAAH.
CC InterPro: IPR001479; Bac-POQ_repeat.
CC InterPro: IPR002372; Bac-POQ_repeat.
CC Pfam: PF01011; Bacterial_POQ_7.
CC PROSITE: PS00363; BACTERIAL_POQ_1; 1.
CC PROSITE: PS00364; BACTERIAL_POQ_2; 1.
CC Oxidoreductase; POQ; Signal; Methanol utilization; Periplasmic.
CC SIGNAL 1 32 METHANOL DEHYDROGENASE SUBUNIT 1.
CC CHAIN 33 631 BY SIMILARITY.
CC DISULFID 135 136 BY SIMILARITY.
CC DISULFID 418 447 BY SIMILARITY.
CC ACT_SITE 335 335 BASE (POTENTIAL).
CC SEQUENCE 631 AA: 69799 MW: 0934DC93FEC5730B CRC64:
SQ
Query Match 15.8%; Score 424; DB 1; Length 631;
Best Local Similarity 24.9%; Pred. No. 3.1e-24;
Matches 153; Conservative 85; Mismatches 201; Indels 176; Gaps 26;
4 TSLMSAGATALLAAPFAOVTPYDEL-LANPPAGESTVSGONENRHSPLTQIT 60
12 SSLMAVAVMLAVLTAPATA-----NDOLVELADPA-NMVTGRDINAQNSEMDIN 65
61 TENVGOLQVW--ARGM-----OPGV--QVTP- 84
66 KENVKQLRPWSTGVHGHGECTPLVGDGMFTPTPTTFLDLNPEGLIMQNKPK 125
85 -----LHSDVMTLANGD-----VIOALDAKGDIMHRRQL 118
126 ONPTARTVACDGVNRLAYW--PGDOOVKPLIFRTQDGHVAMDAETGERM----- 177
119 PNIAITLVANGVYAGTQOYSPF-----GC-----FVSGHDSATGEELMKN 161
178 -----IMNSDKVGSSTLTATVYKIDLVGSGCAELGVAGYVAYDVNSGEMRMA 230
162 YFI-----PRAGEE--GDEYWGNDYEAHMMTGA--WQGITDPVNTLV 200
231 FAIGPEDELLAEDFNAPNPHYGOKNLELEWEGD--AMKIGGTWGWYAYDPEVDLF 287
201 HYGSTVAGRASSTOEGTGTNTFAVAPDTEGELVHMQTLPPDNDQCTEEMV 260
288 YVSGNPAPENETMR--PDKNW--TMAIMGREATGEKFAFYOKTPDEMD-----Y 336
261 TNDVOPSTEMEGLSINPAATGERVYLVGPKTGTGTMGFDAETGEFLWARDNTYOM 320
337 AGVAVWMLSEQEDKO-----GQMRKLLTPDRNGVYITLDRNGDLISA----- 380
321 IESIDE--NGIVTVNEDALIKELDEVY-----DVCEPTFLGGRDPSAALPDGTY 369
381 -DKMDVYVWYKVOQDTGTPVDPDFGRMDHKAADICPSAMGYHNGCHSYDERKVF 439
370 FTLNNVGYDMAVDOFTSMYVNTNSVTKLPPGKM-----IGRTAIDISTGRTIWS 424
440 MGINHICMDWEPMLPYRAGQFVGVATLTPYGPKATLBRAGAGIKAYDAISGEMKE 499
425 VERAANVSPLSTGGGVLENGGTDRYFALSOETGETLMQTLAVASGAQISEVDM 484

```

```

Db 500 KMERFVWGCTAATAGLTFYVTLDFIKARDSDTGLMKFKLPSSVYGHPTKXKDH 559
QY 485 QYVAVGGVSYGSG 499
Db 560 QYVAV-----MYGVG 569
RESULT 9
EXAA_PSEAE STANDARD; PRT: 623 AA.
ID EXAA_PSEAE
AC 092437;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Quinoprotein ethanol dehydrogenase precursor (EC 1.1.99.-) (QEDH).
GN EXAA OR PA1982.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OC NCBI_TaxID=287;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=ATCC 17933;
RC MEDLINE=20437337; PubMed=10984043;
RX Stover C.K., Plam X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,
RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA Gader L.L., Coulter S.N., Tolentino E., Westbrock-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Landig K., Lam R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PAOI, an
RL opportunistic pathogen."
RL Nature 406:959-964(2000).
[3]
RN SEQUENCE OF 1-52 FROM N.A.
RP STRAIN=ATCC 17933;
RC MEDLINE=99173751; PubMed=10075429;
RX Schobert M., Goerlich H.;
RT "Cytochrome c550 is an essential component of the quinoprotein ethanol
RT oxidation system in Pseudomonas aeruginosa: cloning and sequencing of
RT the genes encoding cytochrome c550 and an adjacent acetaldehyde
RL dehydrogenase."
RL Microbiology 145:471-481(1999).
[4]
RN X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF 35-616.
RP STRAIN=ATCC 17933;
RC MEDLINE=20202376; PubMed=10736230;
RX Kettel T., Diehl A., Knaute T., Stezowski J.J., Hoehne W.,
RA Goerlich H.;
RT "X-ray structure of the quinoprotein ethanol dehydrogenase from
RT Pseudomonas aeruginosa: basis of substrate specificity."
RL J. Mol. Biol. 297:961-974(2000).
-1- FUNCTION: OXIDIZE PRIMARY ALCOHOLS AND ALSO ACTS ON SECONDARY
CC ALCOHOL, BUT NOT HIGHLY ACTIVE ON METHANOL.
CC -1- COFACTOR: POQ AND CALCIUM.
CC -1- SUBUNIT: HOMODIMER.
CC -1- SUBCELLULAR LOCATION: Periplasmic.
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POQ DEHYDROGENASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```


CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: AJ009858; CAA08866.1;
 DR EMBL: AE004624; AAC05370.1;
 DR EMBL: AF068264; AAC79657.1;
 DR PDB: 1PLG; 30-AUG-00.
 DR InterPro: IPR001479; Bac_PQQ.
 DR InterPro: IPR002372; Bac_PQQ_repeat.
 DR Pfam: PF01011; Bacterial_PQQ_7.
 DR PROSITE: PS00363; BACTERIAL_PQQ_1; FALSE_NEG.
 DR PROSITE: PS00364; BACTERIAL_PQQ_2; FALSE_NEG.
 DR Oxidoreductase: PQQ; Periplasmic; Signal; Calcium; 3D-structure;
 KW Complete proteome.
 FT SIGNAL
 FT CHAIN 1 35 34
 FT DISUFID 139 140
 SQ SEQUENCE 623 AA; 68123 MW; 32DE5DF20B291D6 CRC64;
 QUNOPROTEIN ETHANOL DEHYDROGENASE.

Query Match 15.4%; Score 413; DB 1; Length 623;
 Best Local Similarity 23.3%; Pred. No. 2,1e-23;
 Matches 140; Conservative 93; Mismatches 245; Indels 122; Gaps 20;

OY 3 PSLIMASGALAL--LAAPAFQVPTDELALN--PPAGEMISGONQENYRSPLT 57
 DB 9 PGLLRPLSLHCLAFVALGSAALADVWEDIANDKTGVLDYGMGTAAORWSPILK 68
 OY 58 QITTEWGOLOLVARGM---OPKQVQVPLHGVMTLANGDVIOAIDKTKPLME 113
 DB 69 QVNAADNVEKLPAPMSYSFGDEKQGO--ESQAIYSDGVIYVYASRSLFALDKTKRLMT 127
 OY 114 HRRLP.N-----IATL--NIVA-----NGVY-----AGST 137
 DB 128 YNHRLPDIPRCDDVNRGAIIYGDKVFETLDASVALNKTGKVKKKFADHAGCYT 187
 OY 138 COYSP-----FCGF--VSGHDSATGEELMRNFYIPR-----166
 DB 188 MTGAPITVKGKGVLLHSGSGDEFGVGRFARDPDTGEIIMRPFEGHNGRLNGK 247
 OY 167 -----AGEGETWGNDEA-----RMMTGA--NGQITYPVNLVHYGSTANGPASET 213
 DB 248 DSTVTGVAAPWPDNRNPTGKVESWMSGGAPWQASFLAETNITITVAGNPGPWNTW 307
 OY 214 ORSTPGG-----TLGVTFRFVRPDGELVNRHQLPRDNMDQCTFFEMATNVADVP 267
 DB 308 APTAKGNPHDYDSLX--TSGOVQVDPSSGEYKWFYQHTPNDAMDFSGNNELVLFDKAD 366
 OY 268 STEMEGLASINP-----AATGERRVLGVPCKTGTWM--QFDAETGEFLMADTYQNM 320
 DB 367 GRIVATATADNNGEYVVDNRNGKLNAPFPVDNITWASHIDLKGT-----RVERREGO 421
 OY 321 IESIDENGITVNEDALIKELDYDVCPTFELGRDMPSSALNPDGSIYFIPLNVVCDYM 380
 DB 422 RPLPPEP-----QKHKAIVEYSPFLDGKNNMNPAYSDGTLEYYVPAHMKDXY 471
 OY 381 MAVQOEFSDMYNTSVNTLPPGKMIGRIDALIDISTGRITMSVERAANSPLYLSGG 440
 DB 472 WTEEVSTTKSAIYGMGFRIKRYDDHVSILRAMDVSQKVVMEKEHPLMAGVLAAG 531
 OY 441 GVLNFGTDYFRFALSQETGETLMOTRLATVASGAISYEDVMQYVAIAGGVSYSGSL 500
 DB 532 NLVFTGTGDYFKAFDAKSGKELMKFQNGSGIVSPITWEDGDGYLGVT---VGYGAV 588

RESULT 10
 DHM1 METME
 ID DHM1 METME STANDARD: PRT; 571 AA.
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)

DE Methanol dehydrogenase subunit 1 (Ec 1.1.99.8) (MDH large alpha
 subunit) (MEDH).
 OS Methylobacillus methylotrophicus (Bacterium W3A1).
 OC Bacteria; Proteobacteria; beta subdivision; Methylobacillus group;
 ON NCBI_Taxid-17;
 RX X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS).
 RA MEDLINE-9405969; PubMed-8241148.
 RA White S., Boyd G., Mathews F.S., Xia Z.-X., Dai W.-W., Zhang Y.-F.,
 RT Davidson V.L.;
 RT "The active site structure of the calcium-containing quinoprotein
 RL methanol dehydrogenase."
 RL Biochemistry 32:12953-12958(1993).
 RP [2]
 RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS).
 RA MEDLINE-93054513; PubMed-1331050.
 RA Xia Z.-X., Dai W.-W., Xiong J.-P., Hao Z.-P., Davidson V.L., White S.,
 RA Mathews F.S.;
 RT "The three-dimensional structures of methanol dehydrogenase from two
 RT methylobacillus bacteria at 2.6-A resolution."
 RL J. Biol. Chem. 267:22289-22297(1992).
 CC -1 CATALYTIC ACTIVITY: A primary alcohol + acceptor -> an aldehyde +
 CC reduced acceptor.
 CC -1 Cofactor: TWO MOLECULES OF PQQ AND TWO MOLECULES OF CALCIUM
 CC PER TETRAMER.
 CC -1 SUBUNIT: THE HOLOENZYME MDH HAS AN ALPHA-2/BETA-2 CONFIGURATION
 CC OF A LARGE ALPHA AND A SMALL BETA SUBUNIT.
 CC -1 SUBCELLULAR LOCATION: PERIPLASMIC, BUT ASSOCIATED WITH INNER
 CC MEMBRANE.
 DR PDB: 4AAH; 08-DEC-96.
 DR InterPro: IPR001479; Bac_PQQ.
 DR InterPro: IPR002372; Bac_PQQ_repeat.
 DR Pfam: PF01011; Bacterial_PQQ_7.
 DR PROSITE: PS00363; BACTERIAL_PQQ_1; FALSE_NEG.
 DR PROSITE: PS00364; BACTERIAL_PQQ_2; 1.
 KW Oxidoreductase; PQQ; Methanol utilization; Periplasmic; 3D-structure;
 KW Calcium.
 FT DISUFID 103 104
 FT ACT_SITE 379 408
 FT ACT_SITE 297 297
 FT HELIX 2 9
 FT TURN 11 12
 FT STRAND 14 14
 FT TURN 17 18
 FT TURN 21 22
 FT STRAND 26 27
 FT TURN 34 36
 FT TURN 37 39
 FT STRAND 41 47
 FT STRAND 59 61
 FT TURN 62 63
 FT STRAND 64 68
 FT TURN 71 73
 FT STRAND 75 79
 FT TURN 80 81
 FT TURN 83 84
 FT STRAND 86 90
 FT HELIX 96 101
 FT TURN 103 104
 FT STRAND 112 114
 FT TURN 115 116
 FT STRAND 117 121
 FT TURN 123 124
 FT STRAND 126 131
 FT TURN 132 134
 FT STRAND 137 142
 FT TURN 146 148
 FT HELIX 149 149
 FT STRAND 151 151
 FT TURN 157 159
 FT TURN 160 161

BASE (POTENTIAL).

Fri May 24 11:27:26 2002

```

FT STRAND 162 166
FT STRAND 168 168
FT STRAND 170 172
FT HELIX 173 173
FT STRAND 177 182
FT STRAND 183 185
FT STRAND 188 193
FT STRAND 198 202
FT STRAND 203 203
FT STRAND 205 210
FT TURN 212 213
FT TURN 219 222
FT HELIX 223 223
FT TURN 226 231
FT HELIX 241 243
FT STRAND 244 247
FT TURN 248 252
FT STRAND 261 263
FT HELIX 269 270
FT TURN 273 278
FT STRAND 279 281
FT TURN 284 289
FT STRAND 305 311
FT STRAND 312 313
FT TURN 314 322
FT STRAND 324 325
FT STRAND 327 332
FT STRAND 333 333
FT TURN 338 343
FT STRAND 346 347
FT STRAND 351 355
FT STRAND 356 359
FT STRAND 360 363
FT STRAND 365 367
FT HELIX 370 372
FT TURN 371 373
FT STRAND 375 378
FT STRAND 382 383
FT STRAND 390 390
FT STRAND 392 394
FT STRAND 395 398
FT TURN 399 405
FT STRAND 407 413
FT STRAND 420 421
FT TURN 427 433
FT STRAND 435 436
FT TURN 438 439
FT STRAND 444 450
FT STRAND 451 454
FT STRAND 455 462
FT STRAND 469 469
FT STRAND 471 473
FT STRAND 474 476
FT TURN 477 481
FT STRAND 483 484
FT TURN 486 491
FT STRAND 492 495
FT TURN 496 502
FT STRAND 512 516
FT STRAND 517 518
FT TURN 519 526
FT STRAND 530 533
FT TURN 534 538
FT HELIX 539 539
FT TURN 543 545
FT TURN 546 552
FT HELIX 553 555
FT TURN 556 558
FT HELIX 559 559
FT STRAND 565 570
FT STRAND 571 AA: 62449 MW: OBE94EA5A2DAB1E1 CRC64;
SEQUENCE

```

```

Query Match 14.4%; Score 388.5; DB 1; Length 571;
Best Local Similarity 23.9%; Pred. No. 1,3e-21;
Matches 136; Conservative 78; Mismatches 199; Indels 155; Gaps 22;

QY 38 AGEWISYGQNEQYRSPILQITTEWGOLOLV--ARGMQPGVQVYPLIHGVMY--- 92
DB 11 AGAMPITAGGYSOHNSPLAQINKSVKVKVKAWSFSTGVINGH-EGAPLVIGDMKTVHS 69
QY 93 -----LANPGDYI----- 129
DB 70 AFPNNNTYALNLDPGKLYMOHKPRKODASTKAVWCDDYDRGLAGQIVKQKANGHLLA 145
QY 103 IDAKTGDLINERHROLNFIATLNIANGVIVAGSTCOISPF-----GC----- 176
DB 130 LDAKTGKINMEVEYCDPRV-----GSTLLOAPFVAKDTVLMGSCGAEIVRG 187
QY 146 FVSGHDSATGEELMR-----NYFIRPAG--EGDETWGNDEYANRMTGA 233
DB 177 AVNAFPLKTGELKWRAPAFATGSDSVRLAKDPSANPHYGFGFGTGTWEGD---AMKIGG 239
QY 188 ---WGQITYPYTNLVHGSTAVGPASETQRTGCG-----TLTGNTFRFAYRPDTGIV 285
DB 234 GTWGWYAYDPKLNLFYSGSNPAPMNETMR--PDGNKWTWTIGRDL-----DTGMAR 297
QY 240 WRHQTLPKRDNDQCTFEAMATNVDOP--STEMEGLIOS-INPNAATGERVVLGVCKTG 326
DB 286 WGYOKTPHDEWDDEAGVNOVLD--QPVNAKMTPLLSHIDRN-----G 326
QY 298 TMMQFDEATGEFIWARDTN-YQNMIESIDENGIVYNEDALIKELDVE-XDVCPTFLGR 355
DB 327 ILYTNRENGNLIYAERKVDPAVNVFKVDKTGTVPYRDEPFRMDHGTINCSANGPH 386
QY 356 DWPSAALNPDGSIYFIPLNVCYDMAVDEFTSMQVYTSVNTLP---PGKDMIGRI 411
DB 387 NOGVDSYDPESKTLIYAGNHICMDPEPMLPYRAGQFEVGAITIAMTPGNSPTKKEMQOI 446
QY 412 DAIDISTGRITLMSVERAANKSPVLSITGGGVLENGGTDYFRALSOETGETIMORLAFV 471
DB 447 RAPDLTFTGKAKWTKEKFAWAGTLYTGGLVWYATLDGYLKALDKRDEKELNRMKPSG 506
QY 472 ASGQATSYEDGCMQVVAIAGGVSYSYG 499
DB 507 GIGSPMTYSFKQKQYI-----GSMWGVG 529

RESULT 11
ID DHG_ECOLI STANDARD; PRT: 796 AA.
AC P15877:
DT 01-APR-1990 (Rel. 14, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update) (EC 1.1.99.17).
DE Glucose dehydrogenase [pyrroloquinoline-quinone] (EC 1.1.99.17).
GN GCD OR B0124.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-K12:
RC MEDLINE-91035240; PubMed-2228962;
RA Cleton-Jansen A.-M., Goosen N., Fayet O., van de Putte P.;
RA "Cloning, mapping, and sequencing of the gene encoding Escherichia
RT coli quinoprotein glucose dehydrogenase.";
RL J. Bacteriol. 172:6308-6315(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-K12 / W3110;
RC MEDLINE-93123180; PubMed-8419307;
RA Yamada M., Asaoka S., Sailer M.H. Jr., Yamada Y.;
RA "Characterization of the gcd gene from Escherichia coli K-12 W3110
RT

```

RT	and regulation of its expression."	
RT	J. Bacteriol. 175:568-571(1993).	
RN	[3]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN-K12 / W3110.	
RA	MEDLINE=94261430; PubMed=8202364;	
RA	Fujita N., Mori H., Yura T., Ishihama A.;	
RT	"Systematic sequencing of the Escherichia coli genome: analysis of	
RT	the 2.4-.1 min (110,917-193,643 bp) region."	
RL	Nucleic Acids Res. 22:1637-1639(1994).	
RN	[4]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN-K12 / MG1655;	
RA	MEDLINE=97426617; PubMed=9278503;	
RA	Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,	
RA	Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,	
RT	Mau B., Shao Y., Kirkpatrick H.A., Goeden M.A., Rose D.J.,	
RT	"The complete genome sequence of Escherichia coli K-12."	
RL	Science 277:1453-1474(1997).	
RN	[5]	
RP	TOPOLOGY.	
RA	MEDLINE=93286127; PubMed=8509415;	
RA	Yanada M., Sumi K., Matsushita K., Adachi O., Yamada Y.;	
RT	"Topological analysis of quinoprotein glucose dehydrogenase of Escherichia	
RT	coli and its ubiquinone-binding site."	
RL	J. Biol. Chem. 268:12812-12817(1993).	
RN	[6]	
RP	3D-STRUCTURE MODELING.	
RA	MEDLINE=96128046; PubMed=8554505;	
RA	Cozier G.E., Anthony C.;	
RT	"Structure of the quinoprotein glucose dehydrogenase of Escherichia	
RT	coli modelled on that of methanol dehydrogenase from Methylobacterium	
RT	extorquens."	
RL	Biochem. J. 312:679-685(1995).	
CC	-1- FUNCTION: GDH IS PROBABLY INVOLVED IN ENERGY CONSERVATION RATHER	
CC	THAN IN SUGAR METABOLISM.	
CC	-1- CATALYTIC ACTIVITY: D-glucose + acceptor = D-glucono-1,5-lactone +	
CC	reduced acceptor.	
CC	-1- COFACTOR: PO.	
CC	-1- SUBUNIT: MONOMER.	
CC	-1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane.	
CC	-1- SIMILARITY: BELONGS TO THE BACTERIAL POO DEHYDROGENASE FAMILY.	
CC	-----	
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration	
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -	
CC	the European Bioinformatics Institute. There are no restrictions on	
CC	use by non-profit institutions as long as its content is in no way	
CC	modified and this statement is not removed. Usage by and for commercial	
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/	
CC	or send an email to license@isb-sib.ch).	
CC	-----	
DR	EMBL; X1333; CA935706.1; -	
DR	EMBL; D12651; BA902174.1; -	
DR	EMBL; D26562; CA820298.1; -	
DR	EMBL; AE000123; AAC73235.1; -	
DR	PIR; JVO107; JVO107.	
DR	HSSP; P38539; 4AAH.	
DR	Ecogene; EG10369; gcd.	
DR	InterPro: IPR001478; Bac_P00.	
DR	InterPro: IPR002372; Bac_P00-repeat.	
DR	Pfam: PF01011; Bacterial_P00_7.	
DR	PROSITE; PS00363; BACTERIAL_P00_1; 1.	
DR	PROSITE; PS00364; BACTERIAL_P00_2; 1.	
KW	Oxidoreductase; POO; Transmembrane; Inner membrane; Periplasmic;	
FT	Complete proteome.	
FT	DOMAIN 1 10	
FT	TRANSXEM 11 37	
FT	DOMAIN 38 40	
FT	TRANSXEM 41 58	
FT	DOMAIN 59 62	
FT	TRANSXEM 63 81	
FT	DOMAIN 82 95	
FT	PERIPLASMIC (PROBABLE).	
FT	CYTOPLASMIC (PROBABLE).	
FT	PROBABLE.	
FT	PERIPLASMIC (PROBABLE).	
FT	PROBABLE.	
FT	CYTOPLASMIC (PROBABLE).	
FT	PROBABLE.	
FT	PERIPLASMIC (PROBABLE).	

Query	6	LIMAS	-----	AGALMALAPAAQVYTVTDELANPPAGEMISTYSONENYRHSPLNQ	58
Db	135	LIMAVGNDPQEIINQTLISADATPAFA	ISVPAQ	-----	DMPAYGRNGEGRSPLNQ 185
Qy	59	ITTEVNGQLOLVMA	-----	ROMQKXV	-QVPLIHGVMYLANGDVIAIDAKTGD 109
Db	186	INADVHNLKEAWFERGVDKQPNDEGITEVPIKGVDTLCTAHQRALDPAASGK	245		
Qy	110	LIMHRRQL	-----		
Db	246	EKMHTDELNTNESFOHVTGRCVSYHEAKAPASPEVMAKCPRIITLPYNDGRILAINAE	305		
Qy	119	-----	PAITLITVAN	-----	GVIVAGS
Db	306	NGKLCTEFANKCVLINQSMNMDPLRGLEYEPTSPITTDKTIIVMAKSYVDNRESTRETSQVI	365		
Qy	148	SGHDSATGELERANVEIPRA	-----	GEEDETMGNDEYAKRMWGANGQUTYPTPTNLY	200
Db	366	RGEVDYNTGELMA	-----	EDGAKDPNALISDEHTEFENS	-----
Qy	201	HGSTAVGPASETORGPBGGLYGTNTREAVRPDPGEIVMHTLPDNDNDQCTREMANV	260		
Db	417	YLPKVTTPDJWGNKRTPEQERYASSI	-LALNATTKLANSYGVHNHDLMD	-----	
Qy	261	TWVDQVSTENEGLOSINPNAATGEBRYLVGPCCTGMMQFAENGEP	-----		
Db	469	--LPAQPT	-----	LADITVN	-----
Qy	311	WARDNTYONNMIESIDNGI	-----	VTVNDALIKEL	-DVEYD
Db	519	GAAGKGVYTPLOPFSLSRPYTKDLGADMMGATMEDOLVCVMEFHQNKYBEITFPSEQ	578		
Qy	347	-----	VCPTFLGGRDPSALNP	-----	
Db	579	GTLPVFGNIGMEWGISVDNPREVALIANMALPEYSKILPPGCPNPEQPKAKGTGE	638		
Qy	366	SGT	-----	YPIPLANCYDMAVODEFISMVDYNTSWTKLPPGKMIGRIDALISY	418
Db	639	SGTQPOYGVPGVTLN	-----	PLS	-----
Qy	419	GRTLMSVERAANAS	-----	PV	-----
Db	680	NEVYMKRIRIGIPQMSFEPPEVPPVBNPMGMPLGIPSTAGCNVLFIAATADNITLRAVMS	739		
Qy	459	TGELIMORLATAVASGOA	-----	ISYEVDGMQYVAI	-AGGVSVGS 498
Db	740	NGEKLMOGRLP	-----	AGQATPMTYEVNKGQYVVISAGHGSGFT	780

DE (EC 1.1.99.17) (Quinoprotein glucose DH) (GDH-A).
 GN GDH.
 OS Acinetobacter calcoaceticus.
 OC Bacteria; Proteobacteria; gamma subdivision; Moraxellaceae;
 OC Acinetobacter.
 OX NCBI_TaxID=471;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-LMD 79.41;
 RX MEDLINE=8828368; PubMed=3399393;
 RA Cleton-Jansen A.-M., Goosen N., Odle G., van de Putte P.;
 RT "Nucleotide sequence of the gene coding for quinoprotein glucose
 dehydrogenase from Acinetobacter calcoaceticus."
 RT Nucleic Acids Res. 16:6228-6228(1988).
 RL
 CC -1- CATALYTIC ACTIVITY: D-glucose + acceptor = D-glucono-1,5-lactone +
 reduced acceptor.
 CC -1- COFACTOR: POQ.
 CC -1- SUBUNIT: MONOMER.
 CC -1- SUBCELLULAR LOCATION: Periplasmic.
 CC -1- MISCELLANEOUS: ACINETOBACTER CALCOACETICUS CONTAINS TWO DIFFERENT
 POQ DEPENDENT ENZYMES WITH GDH ACTIVITY. GDH-A PREFERENCES
 2-DEOXYGLUCOSE AS SUBSTRATE, THE SPECIFIC SUBSTRATES FOR GDH-B ARE
 DISACCHARIDES.
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POQ DEHYDROGENASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@sib-sib.ch).
 CC
 CC EMBL: X0735; CA30222.1;
 DR PIR: S00943; S00943.
 DR InterPro: IPR001479; Bac_POQ.
 DR InterPro: IPR002372; Bac_POQ_repeat.
 DR Pfam: PF01011; Bacterial_POQ_7.
 DR PROSITE: PS00363; BACTERIAL_POQ_1; 1.
 DR PROSITE: PS00364; BACTERIAL_POQ_2; 1.
 KW Oxidoreductase; POQ; Transmembrane; Periplasmic; Signal.
 FT SIGNAL 1 33
 FT CHAIN 34 801
 FT TRANSMEM 39 55
 FT TRANSMEM 59 79
 FT TRANSMEM 94 108
 FT TRANSMEM 119 138
 FT ACT_SITE 471 471
 FT SEQUENCE 801 AA: 86956 MW: 2FAEA322EB5450D8 CRC64;
 SO
 Query Match 11.7%; Score 315; DB 1; Length 801;
 Best Local Similarity 21.2%; Pred. No. 6.4e-16;
 Matches 148; Conservative 76; Mismatches 192; Indels 282; Gaps 30;
 QY 20 PAFAGVTPTDELLANPAGEWISYQGNQENYHSPITQITTEVNCQDOLVWARGHOGK 79
 DB 152 PETAADVGVAF-----SDMPAYGRTQAGVRSPLKQINDQNVLDLVAMT--LRIGD 202
 QY 80 V-----QVPLIHGVNMIANDGVIOIDAKTKDGLIWEHRRL----- 118
 DB 203 LKINDSGETVQVPIKIGNNMFLCTAHQQLAIDPATGKEKWRFPDKTKDSFOHLT 262
 QY 119 ----- 118
 DB 263 CRGVWYDANNTTEFATSLQSKSSSTQCPKRVFVNVNDGLVAVNADTKACTDGGNG 322
 QY 119 -----PNATLNIYANGVIAGS--TCQYS--PFGCEVSGHDSAT 154
 DB 323 QVNIQEDMPYAPGYNPISPGIYV-----GSHVVIAGSVTNIYSNKEPSG--VIRGIVYNT 377
 QY 155 GEELMRNYFIPRAGE-----EGDETGWNDYARMTATGMOQIYDPTNLVHYGSTAVG 208

DB 378 KGLIW--VEDTGAADPNNAHGECTTFVHNS-----PNAAPLAYDAKLDIV-YVPTGV- 427
 QY 209 PASERQGTGGTLYG-----TNTFAVAPPDGELVWFHQTLPDNDQDECTFEM 258
 DB 428 -----GNP--DIWGGRTELKERYANSMALINASTGLKVNPFQTHHDMQMDVPSOP 478
 QY 259 MTYNDVDPSTMEGLQSLNP--NNAATGERRVLTGPV-----CKTG--TMMQF 302
 DB 479 SLADINKKAGQTPATVYILTKGNAFVLDNR--NGQPIVPTKRPQVTKRPQTKGEF 536
 QY 303 DAETGEF-----LMA-----RDTNYOMIESIDENGIVTNE 334
 DB 537 YSKTQPSDLNLAPODKLTKDMGATMLDLMCAVSKRLNIGITTPSENCTL----- 592
 QY 335 DALKEIDVYDVCPTFLGGRDPMASALNPS-----GITYF----- 371
 DB 593 -----VPGNLGVFEWGMGMSVNPDRQVAVMNPIGLFPVSRLLIPDPNRAQTK 640
 QY 372 -----PLNNVCYDMAVDOEFTSMQVNTSVTKLPKGDMDIGRIDADISTGRTL 422
 DB 641 GAGTEGGVQPMYGVVY-----GVEISAFSLPGLPKQPMAGVYAGVDLTKHREV 690
 QY 423 W-----SVERAANYSPVLSTGGGVLFNGST--DRYFRALSOETGEL 463
 DB 691 WKRRIGTIRDSLPLNLPOLPAVKIGVPGIAGSISTAGVMMFVGFQDNYLRATVNTNGKRL 750
 QY 464 MQTRLATVAGQA--ISTEVDGMQYVAL-AGGGVSYGS 498
 DB 751 WEARLP--AGGQATPMYIEINGKOVYVIMAGHGSFOT 786
 RESULT 13
 DHG_GLUOX STANDARD: PRT: 808 AA.
 ID DHG_GLUOX
 AC P27175;
 DT 01-AUG-1992 (rel. 23, Created)
 DT 01-AUG-1992 (rel. 23, Last sequence update)
 DT 16-OCT-2001 (rel. 40, Last annotation update)
 DE Glucose dehydrogenase [pyrroloquinoline-quinone] precursor
 DE (EC 1.1.99.17).
 GN GDH.
 OS Gluconobacter oxydans (Gluconobacter suboxydans).
 OC Bacteria; Proteobacteria; alpha subdivision; Acetobacteraceae;
 OC Gluconobacter.
 OX NCBI_TaxID=442;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92017653; PubMed=1833618;
 RA Cleton-Jansen A.-M., Dekker S., van de Putte P., Goosen N.;
 RT "A single amino acid substitution changes the substrate specificity
 of quinoprotein glucose dehydrogenase in Gluconobacter oxydans."
 RT Mol. Gen. Genet. 229:206-212(1991).
 RL [2]
 RN REVISION TO 213.
 RA Goosen N.;
 RA Submitted (FEB-1992) to the EMBL/GenBank/DBJ databases.
 RU
 CC -1- CATALYTIC ACTIVITY: D-glucose + acceptor = D-glucono-1,5-lactone +
 reduced acceptor.
 CC -1- COFACTOR: POQ.
 CC -1- SUBCELLULAR LOCATION: OUTSIDE OF THE PERIPLASMIC MEMBRANE.
 CC -1- MISCELLANEOUS: THE P1 FORM CAN OXIDIZE ONLY D-GLUCOSE, WHILE THE
 P2 FORM CAN ALSO OXIDIZE MALTOSE. THE SEQUENCE OF P1 FORM IS SHOWN
 HERE.
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POQ DEHYDROGENASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@sib-sib.ch).
 CC

CC EMBL: X62710; CAA44594.1; ALT_SEQ.
 DR PIR: S17716; OPKEX.
 DR InterPro: IPR001479; Bac_PQQ.
 DR Pfam: PF01011; Bacterial_PQQ_7.
 DR PROSITE: PS00363; BACTERIAL_PQQ_1; 1.
 DR PROSITE: PS00364; BACTERIAL_PQQ_2; 1.
 DR Oxidoreductase; PQQ; Transmembrane; Periplasmic; Signal.
 FT SIGNAL 1 33
 FT CHAIN 34 808
 FT TRANSMEM 35 54
 FT TRANSMEM 59 76
 FT TRANSMEM 94 108
 FT TRANSMEM 123 138
 FT ACT_SITE 470 470
 FT VARIANT 788 788
 SQ SEQUENCE 808 AA; 87567 MW; 0F4160DA78652445 CRC64;

Query Match
 Best Local Similarity 10.7%; Score 287.5; DB 1; Length 808;
 Matches 140; Conservative 82; Mismatches 229; Indels 251; Gaps 25;

QY 13 ALALLAFAFQVPTVDLLA-----NPRAGEWISYGNQENYRSHPLQIT 61
 DB 128 AVLAFAFLFTDPHDSGLPTQANSPADPVPSEHAYGRQAGDRMSPLNQINA 187
 QY 62 ENVGLOLVW-----ARGNQEKV--QVPLIHGVNVLNPGVQIDAQKGLI 112
 DB 188 TNSVSLAKAMHITKMMNSNDPEQTEATPIEFNNITLQMSLHOKLFAVDGATGNVK 247
 QY 113 EHRQL--PIATLNI-----VANGVIY----- 133
 DB 248 VYDKLQINPGFOLHCRGVSEFEPANAMDSGNAPAPDCAKSDILPVDNGRLVEDAD 307
 QY 134 AGSTC-----QSPGFCV----- 147
 DB 308 TGTGSGFGNNGEIDLRVNPQPTTGGVETSPVITDKLIANSATIDGSKVAKSGA 367
 QY 148 -SGHDSATGEELW---RNYETPRAGEEGDEFTWGDYDARMTGAMQIITDPTNLVHY 202
 DB 368 TQAFDYVTKRWVFDASNDPNQLPDESHPEHNSNSNIVS---SYDANLNLVYI 422
 QY 203 GSTAVPASETORPT---GGLYGTNRF---AVRPDTGELVNRHQTLPDNDQEC 254
 DB 423 PMGV-----GTPDMGDKTKDSERFAGIVALNADTGKLAWEYQTVHHDLMDEL 473
 QY 255 TREMAVTVNDVQPTSEMEGLQSLNPNMATGERRVLTGVPCKTGMNQFAETG-EFLMAR 313
 DB 474 PSQPSLVYDQDGLVPAIYA-----PRTGDIYVLDKRTGKEIVAP 517
 QY 314 DT-----NYQNMIESIDENGIVTVNEDALIKELDV-----EAD 346
 DB 518 ETPVPOGAPGDHSPQPMQ--LTLRPKNPLNSDWTGGITPDQECSTYFTLATE 574
 QY 347 -----VPTFLGRDMPASALNDGSIYIPINNYCYDMAVQD----- 385
 DB 575 GPTTPPSIKSLIPGLGFMWGLAVPQOVAFANPISLPEVSQVLPNGPNLWPE 634
 QY 386 -----EFTSMVYNTSNTKLPQKDMT-----GRIDAIDISTGRTL 422
 DB 635 ENKGTGGEGTGLQHNIGIPYAVNLHFDLPVLPFGIMPKRTPPWGVAGIDTKTKVY 694
 QY 423 WSEVERAANS-----PV-----LSTGGVLF--NGGDRFRALSGETG 460
 DB 695 WQHRNGLTLDMSKGLSLPIPLPIKIGVPSLGLPSTAGNLGFLTAMDYIYIAYNLITG 754
 QY 461 ETLMTQRLTAVASGAISYEVDGMQYVAIAGVSYSGSLNS 502
 DB 755 KYLMODRLPAGAGATPTIYAIANGKQYI-----VTAGGHS 790

RESULT 14
 ID OUIA.ACICA STANDARD; PRT; 809 AA.
 AC 059086;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Quinate/shikimate dehydrogenase [Pyroloquinoline-quinone]
 DE (EC 1.1.99.25) (NAD(P)-independent quinate dehydrogenase);
 GN OUIA.
 OS Acinetobacter calcoaceticus.
 OC Bacteria; Proteobacteria; gamma subdivision; Moraxellaceae;
 OC Acinetobacter.
 OX NCBI_TaxID=471;
 RN [1]
 RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
 RC STRAIN-BD413 / ADP1.
 RX MEDLINE=95095936; PubMed=8002591;
 RA Elsemore D.A., Ornston L.N.;
 RT "The pca-pob supraoperonic cluster of Acinetobacter calcoaceticus
 RT contains quia, the structural gene for quinate-shikimate
 RT dehydrogenase.";
 RL J. Bacteriol. 176:7659-7666(1994).
 RN [2]
 RP SEQUENCE OF 1-20 FROM N.A.
 RC STRAIN-BD413 / ADP1.
 RX MEDLINE=96011389; PubMed=7592351;
 RA Elsemore D.A., Ornston L.N.;
 RT "Unusual ancestry of dehydrogenases associated with quinate catabolism
 RT J. Bacteriol. 177:5971-5978(1995).
 CC -1- FUNCTION: CAN ACT EITHER ON QUINATE OR ON SHIKIMATE.
 CC -1- CATALYTIC ACTIVITY: Quinate + pyrroloquinoline-quinone = 5-
 CC -1- dehydroquininate + reduced pyrroloquinoline-quinone.
 CC -1- CATALYTIC ACTIVITY: SHIKIMATE + PYRROLOQUINOLINE-QUINONE = 3-
 CC -1- dehydroshikimate + REDUCED PYRROLOQUINOLINE-QUINONE.
 CC -1- COFACTOR: PQQ.
 CC -1- PATHWAY: QUINIC ACID CATABOLIC PATHWAY: FIRST STEP. THIS PATHWAY
 CC ALLIQU GROWTH OF BACTERIA WITH QUINATE BY ITS CONVERSION TO
 CC PROTOCATECHUATE AND SUBSEQUENT METABOLISM BY THE BETA-KETOADIPATE
 CC PATHWAY.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -1- INDUCTION: BY PROTOCATECHUATE.
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL PQQ DEHYDROGENASE FAMILY.
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>
 CC or send an email to license@sdb-sib.ch).
 CC EMBL: L05770; AAC37161.1; -
 DR InterPro: IPR001479; Bac_PQQ.
 DR Pfam: PF01011; Bacterial_PQQ_7.
 DR PROSITE: PS00363; BACTERIAL_PQQ_1; FALSE NEG.
 DR PROSITE: PS00364; BACTERIAL_PQQ_2; FALSE NEG.
 KW Oxidoreductase; PQQ; Quinate metabolism; Transmembrane.
 FT TRANSMEM 14 34
 FT TRANSMEM 41 61
 FT TRANSMEM 68 88
 FT TRANSMEM 90 110
 FT TRANSMEM 127 147
 FT TRANSMEM 147 147
 SQ SEQUENCE 809 AA; 88196 MW; 71F67CEBFA62BFCB CRC64;

Query Match
 Best Local Similarity 9.3%; Score 250.5; DB 1; Length 809;
 Matches 144; Conservative 70; Mismatches 242; Indels 223; Gaps 24;

Fri May 24 11:27:26 2002

walick-934-125.pwp.rsp

Page 14

```

QY 415 DISTGTLMSVERAAN-----YSPV-----LSTGGGV.FPNCGT-DRYERA 454
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 670 DLRKRSIAWVPEVGFYQDTPGFGIKMLPLPIGMPPLGTLSTGGGLVFIFIGTQDYLLRA 729
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 455 LSGDTEGTLMSQRRLATVASGAISY--EVDGMVYAIAGG 493
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 730 FDSATCKELMKRGLPVGSGGQGPITVYSHKTKGQKQYVVIASGG 770
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

```

Search completed: May 24, 2002, 10:30:52
Job time: 630 sec

Fri May 24 11:27:28 2002

walick-934-125.psp.rsp

Page 1

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2002, 10:30:15 ; Search time 93.8 Seconds
(without alignments)
925,836 Million cell updates/sec

Title: WALICK-934-125.PEP
Perfect score: 2689
Sequence: 1 MKPTSLMASAGALALAAP.....GMQVYAGGVSYSGILNS 502

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 17294929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: SP archaea: *
2: SP bacteria: *
3: SP fungi: *
4: SP human: *
5: SP invertebrate: *
6: SP mammal: *
7: SP mhc: *
8: SP ornamental: *
9: SP phage: *
10: SP plant: *
11: SP rodent: *
12: SP virus: *
13: SP vertebrate: *
14: SP unclassified: *
15: SP virus: *
16: SP bacteriophage: *
17: SP archaea: *

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1490	55.4	608	2	093RE9
2	520.5	19.4	698	2	09KH03
3	487	18.1	742	2	053362
4	482	17.9	601	16	09EYW8
5	481	17.9	601	16	09EYW8
6	479.5	17.8	708	2	046444
7	477.5	17.5	629	2	09E902
8	470.5	17.2	601	2	09A048
9	463.5	17.2	601	2	09A595
10	461.5	17.2	601	2	09A595
11	431.5	16.0	633	2	024759
12	428	15.9	623	2	09AGW3
13	419	15.6	695	2	091935
14	404.5	15.0	695	2	093460
15	391.5	14.6	573	2	059540
16	297.5	11.1	790	2	09X255

Result	1	PRELIMINARY	PRT	608 AA
AC	093RE9	093RE9		
DT	01-DEC-2001 (TREMBLrel. 19, Created)			
DT	01-DEC-2001 (TREMBLrel. 19, Last sequence update)			
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)			
DE	ALCOHOL DEHYDROGENASE.			
GN	ADH.			
OS	Pseudoglucobacter saccharotogenes.			
OC	Bacteria: Pseudoglucobacter.			
OX	NCBI TaxID-133921.			
RN	SEQUENCE FROM N.A.			
RP	SRRAIN-1FO 1464.			
RC	Sribata T., Saito Y.			
RA	"Alcohol dehydrogenase."			
RT	Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.			
RL	EMBL: AB046580; BAB62258.1; "			
DR	EMBL: AB046580; BAB62258.1; "			
SQ	SEQUENCE 608 AA; 65101 MW; 0ACEC97AE11BA570 CRC64;			

ALIGNMENTS

Query Match	Score 1490; DB 2; Length 608;
Best Local Similarity 50.0%; Pred. No. 6,6e-101;	
Matches 291; Conservative 63; Mismatches 138; Indels 90; Gaps 6;	
QY	10 SGALALIAAPAFQ-----VTPVDLLANPAGEWISYQNDEN 50
DB	16 STALIASGPFAPFDHANAAPERSKAGQSAIENFOVTDADLAKNPAMPTLRNGYOG 75
QY	51 YRHSPLTQITTEVNGQDLVWARGMOPGVOTPLIDGVYLANPGDVIAIDATKTDGL 110
DB	76 WGVSPDQINKNVGDLQVLSRTMEPSNGAIAVNGVIFLGNTDNDVDAIDKGTSL 135
QY	111 IWEHRQOLN-----IATLN----- 125
DB	136 IWEHRQOLN-----IATLN----- 174
QY	126 -----IVANGVYAGSTGCTGSPGCTVSGHDSATGELMRNYFIPRAEBDEET 174

```

Db 196 VEEGVANSSGPIVNDGVVIASTQCFSGCYVGTDAESGEELMNTFIPRGEEDDT 255
175 WGN-DYKRWMTGANGQITTPYVNLVHYSTAGPASEFORCPGILYGTNTREAVP 233
256 WCGAFYERMMWTGAMGQITTPYVNLVHYSTAGPASEFORCPGILYGTNTREAVP 315
234 DTGEIWRHQTLPDRNDQDECFEMMYVNDVOPSTEMEGLOSINPMNATGE-RVLTGY 292
316 KTGEVYWKHQTLPDRNDQDECFEMMYVNDVOPSTEMEGLOSINPMNATGE-RVLTGY 375
293 PCKTGMOPDAETGEFLMARDNTYOMIESIDENGIVVNDALTKLELVEYDVCPTFL 352
376 PCKTGMOPDAETGEFLMARDNTYOMIESIDENGIVVNDALTKLELVEYDVCPTFL 435
353 GGRDWPSSALNPDGSIYFPLNNVCYDMMAYDOESTMDVYNTSNVTKLPCKDMIGRID 412
436 GGRDWPSSALNPDGSIYFPLNNVCYDMMAYDOESTMDVYNTSNVTKLPCKDMIGRID 495
413 AIDISGRTLMSEVERAANYSPVLTSGGVLFGNGTDRYFALSOETGELMGTRLATYA 472
496 AIDALGETKMSYETRLALYDPLVLTGGDLVFGGIDRFDALDAESGEVWSTRLPGAV 555
473 SGQALSYEDGMQYVAIAGG-----VYSGSLN 501
556 SGTTSYSIDGRQYVAVVSGSLGPTFGPTTDPVDSASGAN 597

RESULT 2
09KH03 PRELIMINARY; PRT; 698 AA.
AC 09KH03;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DE TERRAHYDROFURYL ALCOHOL DEHYDROGENASE
OS Alkaligenes eutrophus (Ralstonia eutropha)
OC Bacteria; Proteobacteria; beta subdivision; Ralstonia group;
OC Ralstonia.
OC NCBI_TaxID=510;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-21122557; PubMed-11222593;
RA Zarit G., Schrader T., Andreesen J.R.;
RT Tetrahydrofuryl Alcohol Dehydrogenase from Ralstonia eutropha
RT Tetrahydrofuryl Alcohol Dehydrogenase from Ralstonia eutropha
RL U. Bacteriol. 183:1954-1960(2001).
DR EMBL; AF277373; AF86335.1;
DR HSSP; Q924J7; IFLG.
DR InterPro; IPR002372; Bac_P00_repeat.
DR InterPro; IPR003045; CytC_heme_bind.
DR Pfam; PF01011; Bacterial_P00; 6.
DR Pfam; PF00034; Cytochrome_c; 1.
DR PROSITE; PS00190; CYTOCHROME_C; UNKNOWN.1.
DR PROSITE; PS00190; CYTOCHROME_C; UNKNOWN.1.
SQ SEQUENCE 698 AA; 75471 MW; 9C87F180BD1D2FB2 CRC64;

```

Query Match 19.4%; Score 520.5; DB 2; Length 698;
 Best Local Similarity 27.4%; Pred. No. 9e-30;
 Matches 160; Conservative 80; Mismatches 216; Indels 129; Gaps 21;

```

OY 10 SAGALALAAPFA--QVTPVDELALNPAG--EMISYCONDENRHSLTQITTEVNG 65
Db 14 AAASVALPAPAFANAAARVDGAIRANEGTPNMPSTGLDYAEIRFSLQEVNAGNR 73
OY 66 QLOLVWAGMOPGK-VQVTPLIHDGVYLANPGDVIOAIDAKTGDLIMHRRQLP----- 119
Db 74 NIGLAWSTDLSTGVEATPLVVDGVVYVAPSVVHAIDAKTGKRLMWTYDQVPRDQAY 133
OY 120 -----NIATLNTIVA 128

```

```

Db 134 KCCDDVNNNGVALYOGKVFAGAFDRLVAIDAAGKKVMEQDITVDRSKSYTTTGAPRY 193
129 NG-VIVAGSTQCYSPGCFVSGHDSATGEELMNTFIP-----RAGEGDETFWGD 178
194 NGKVLINGCAIYVGRG-YITAYIDETGKQGMRYVTPGPAPFENEAAMAKAAAT--D 250
OY 179 YEAR-WMTG-----AMGQITTPYVNLVHYSTAGPASEFORCPGILYGTNTREAVR 232
251 PCKTGMOPDAETGEFLMARDNTYOMIESIDENGIVVNDALTKLELVEYDVCPTFL 309
310 PCKTGMOPDAETGEFLMARDNTYOMIESIDENGIVVNDALTKLELVEYDVCPTFL 351
291 GVPCKTGMOPDAETGEFLMARDNTYOMIESIDENGIVVNDALTKLELVEYDVCPTFL 348
352 HAP-KNGFFVYIDRINKGFLISAKNFVDVWMAAGYDKNGRPVETPOADISGRPA---DVV 406
349 PTFGLGHDWPSALNPDGSIYFPLNNVCYDMMAYDOESTMDVYNTSNVTKLPCKDMIGRID 461
407 GGPFGAHNMHSMFHPKLGIAFLTPAQHVPL-TLADNKM-----VHNOKDSEPAHRYGN 497
398 -----VTKLPCKDMIGRIDAIDISTGRTLMSEVERAANYSPVLTSGGVLFGNGTDRYF 453
462 LGMLVNAEPPRSKPMGRILAMPDLAOKAVWHRHAGFMNGGTLATAGNLVFGGTADGRV 521
OY 454 ALSOETGELMGTRLATYASGOAISYEVDGMQYVAIAGG-----VYSGSLN 501
522 AYHATGELMGTRLATYASGOAISYEVDGMQYVAIAGG-----VYSGSLN 597

```

```

RESULT 3
053362 PRELIMINARY; PRT; 742 AA.
AC 053362;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DE DEHYDROGENASE SUBUNIT OF ALCOHOL DEHYDROGENASE
OS Acetobacter pasteurianus (Acetobacter turbidans)
OC Bacteria; Proteobacteria; alpha subdivision; Acetobacteraceae;
OC Acetobacter.
OC NCBI_TaxID=438;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=NC11380;
RX MEDLINE-94042848; PubMed-8226628;
RA Takemura H., Kondo K., Horinouchi S., Beppu T.;
RT Induction by ethanol of alcohol dehydrogenase activity in Acetobacter
RT pasteurianus.
RT U. Bacteriol. 175:6857-6866(1993).
DR EMBL; D13893; BAA40252.1;
DR HSSP; Q924J7; IFLG.
DR InterPro; IPR001479; Bac_P00.
DR InterPro; IPR002372; Bac_P00_repeat.
DR Pfam; PF01011; Bacterial_P00; 6.
DR PROSITE; PS00363; BACTERIAL_P00; 1.
DR PROSITE; PS00363; BACTERIAL_P00; 2.
SQ SEQUENCE 742 AA; 81680 MW; 031C4E9B6B13EC98 CRC64;

```

Query Match 18.1%; Score 487; DB 2; Length 742;
 Best Local Similarity 26.5%; Pred. No. 2.8e-27;
 Matches 161; Conservative 81; Mismatches 215; Indels 150; Gaps 21;

```

OY 11 AGALALAAPFAQVTPVDELALNPAG--EMISYCONDENRHSLTQITTEVNG 68
Db 23 AALPVAAPARADQAGTGEALIHADHREMTSTGRTYSEGRYSLDINSNGDGLK 82
OY 69 LWARGMOPGKQ-VQVTPLIHDGVYLANPGDVIOAIDAKTGDLIMHRRQLP-NIA----- 122
Db 83 LAMYYTLDITNGQKATPLVVDGVVYVAPSVVHAIDAKTGKRLMWTYDQVPRDQAY 142

```

```

QY 123 --TLN-----I 126
DB 143 CQTVNRGAGYNGKVGWCTGFDGLVAADAKTGKWEVNTIPADASTGKORSYVDGAVR 202
QY 127 VANGYIYASTCOYSPEGC--FVSGHDSATGEELMNTYFIPRAGEGD-----172
DB 203 VAKGLVILGN--GSEFGARGFVSFADETGKLMREYIVPNKNEPDHADVNTLMSKA 260
QY 173 -ETWMDYEARMWT-----GAMQITDPTVNLVHGSTAVPASERGTGPGTLY---223
DB 261 YKTMGP--KGAVYRGGGGYVWDSLVYDPSDLX--LAVG-----NCSPPNRYRSE 309
QY 224 --GTN-----TRFVAPDGLVWRHQTLPDNDQECTFEKATVNVQVSTEGLOST 277
DB 310 GIGSNLFLGSIYALKRGEYVWHFOATPMQOMYTSVQOIMTLDMV-----357
QY 278 NPNAAATGE--RVILGVPCKTGTMQFDATGEFLWADPTNQNMIESTIDE--NGIIVYND 335
DB 358 -----NEMRHVIMHAP--KNGEFTYLDATGTEFLAGKNVYVQNNANGLDPLTGPIYND 411
QY 336 AILKELDERVDCPTFLGGRDPSAALNDPSGIYFIPLNNVY-----DMA 382
DB 412 GLYTLTKFVWIGPLGAGNFMGMAYSPTKTLVYIPAHQIPFGYKNQYGGFKHPDAMN 471
QY 383 VDQFTSMQVNTNSNTKLPKPKDMIGRIDADISTGRTLSVERAANYSPVLTGGGV 442
DB 472 VGLDMTKNGLDPTPE--ARTAYIKDLHGMILLAMPVKMEYVWKIDHKGPMNGSVLTGDL 530
QY 443 LFNQSTDRYFALSOETGELTMTRELATVASCQAISEYDGMQYVA-----IAG 492
DB 531 LFQGLANEFHAYDATNGSDLYKFDQSGIILASPMYISVNGKYVAVEYGMGIVPISMG 590
QY 493 GYSYSG 499
DB 591 GVGRTSG 597

```

RESULT 4

Q9EYV8 PRELIMINARY; PRT; 601 AA.

AC 09EYV8; 16, Created)

DT 01-MAR-2001 (TEMBLrel. 16, last sequence update)

DT 01-DEC-2001 (TEMBLrel. 19, last annotation update)

DE 01-DEC-2001 (TEMBLrel. 19, last annotation update)

DE METHANOL DEHYDROGENASE LARGE SUBUNIT-LIKE PROTEIN.

GN "KAP"

OS Rhizobium meliloti (Sinorhizobium meliloti).

OS Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;

OC Rhizobiaceae; Sinorhizobium.

OC NCBI_Taxid=382;

OX NCBI_Taxid=382;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=RM1021;

RA Penner B.J., Tiwari R.P., Dilworth M.J.;

RT "Regulation of C1 assimilation in Sinorhizobium meliloti."

RT Submitted (SEP-2000) to the EMBL/GenBank/DBD databases.

RL EMBL: AF309488; AAC31643.1; "

DR HSP: P38539; 4AAH

DR InterPro: IPR002372; Bac_PQQ-repeat.

DR Pfam: PF01011; Bacterial_PQQ_7

DR SEQUENCE 601 AA; 6586 MW; 09F743726B2E2A83 CRC64;

Query Match 17.9%; Score 482; DB 2; Length 601;

Best Local Similarity 25.8%; Pred. No. 4; de-27;

Matches 155; Conservative 90; Mismatches 206; Indels 150; Gaps 22;

QY 14 LALLAFAFAFQVTPVDEL--LANPAGWISYQONENYHSPLOTITENGGQLDLY 71

DB 8 LAIWSGGGAVAFANDELQKLDDP--NOMALQOTGDYANLKRSLQDINKDNVGLQYAV 66

QY 72 --ARGMQPGKVOVPLIHGCVNTLANP-----GVY-----99

```

DB 67 TFSGLVLRGH--EGSPLVIGDLMTVYTPPNTVVALDLSKDGQVWKTEPKODPNVLPVNC 125
QY 100 -----IADAKTGDLI-----EHRROLPTATINIVANGY 131
DB 126 CQTVNRGAGYVADNKKIFLQADTTVALADKTKGKATWSKKNQNDATKGETNATVMPYDKI 185
QY 132 IVAGSYCOYSPEGCFSVSGHDSATGEELMNTYF-----PRAGEGD 173
DB 186 LVGISGGEVRG--HTATA SMABGVLMRGYSMGPSDYLIDPEKTYHLKRPVGRKDSGLT 244
QY 174 TWGMDYEARMWT-----AMQITDPTVNLVHGSTAVPASERGTGPGTLYGNTTFA 230
DB 245 TWEGD--QWKIGGTTWGYSTDEPNLYTGTGNPTWNPOTR--FGDNRW--SMTIFA 298
QY 231 VRPTGELVWRHQTLPDNDQECTFEKATVNVQVSTEGLOSTINPNATGERVLT 290
DB 299 RDVDTGAKKILYQMTPHDEMDYGVNEMILTGQHDGK-----DRKLLT 342
QY 291 GVPCKTGMQFDATGEFL-----WARDVYQNMIE-----SIDENGI 329
DB 343 HED--RNGFGYTMDEVTELLVAEKYDPTVWATVEYMDKSKYGRPYVAYOSTEONG- 400
QY 330 VVVEDALILKELDERVDCPTFLGGRDPSAALNDPSGIYFIPLNNVYDMAVDOEFTS 389
DB 401 -----EDTNT-----GCPALGKTKDOOPAAVSPKTELYVFTNHVCHDYEPRYSYA 450
QY 390 MDVNTSNVTKLPKPKDM--IGRIDADISTGRTLSVERAANYSPVLTGGGV 446
DB 451 GQYVGAATLSMP--KDSHGAGNFIACDNKEGKIKMSLPESVMSGALATAGDVYFG 509
QY 447 GMDRYFALSOETGELTMTRELATVASCQAISEYDGMQYVAING-----GYSYSG 501
DB 510 TLEGIKAVDAATGKELYRFTKPSGVIGNVMTYAREGQYAVVLSGVGNAGIGLALGT 569
QY 502 S 502
DB 570 N 570

```

RESULT 5

Q9ZMY9 PRELIMINARY; PRT; 601 AA.

AC 09ZMY9; 19, Created)

DT 01-DEC-2001 (TEMBLrel. 19, last sequence update)

DT 01-DEC-2001 (TEMBLrel. 19, last annotation update)

DE 01-DEC-2001 (TEMBLrel. 19, last annotation update)

DE PUTATIVE METHANOL DEHYDROGENASE PROTEIN, LARGE SUBUNIT (EC 1.1.99.8).

GN Smb20173.

OS Rhizobium meliloti (Sinorhizobium meliloti).

OS Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;

OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;

OC Rhizobiaceae; Sinorhizobium.

OC NCBI_Taxid=382;

OX NCBI_Taxid=382;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=1021;

RA MEDLINE=21396508; PubMed=11481431;

RA Finan T.M., Weldner S., Wong K., Bunmester J., Chain P., Gouzy J., Vorhoelter F.J., Hernandez-Lucas I., Becker A., Cowie A., Gouzy J.,

RA "The complete sequence of the 1.683-kb PSYMB megaplasmid from the N2-

RT "fixing endosymbiont Sinorhizobium meliloti."

RT Proc. Natl. Acad. Sci. U.S.A. 98:9889-9894(2001).

RL EMBL: AF603642; CAC48573.1; "

DR Oxidoreductase; Plasmid; Hypothetical protein; Complete proteome.

DR SEQUENCE 601 AA; 65759 MW; D73424EFBD15ADB8 CRC64;

Query Match 17.9%; Score 481; DB 16; Length 601;

Best Local Similarity 25.8%; Pred. No. 5; 7e-27;

Matches 155; Conservative 90; Mismatches 206; Indels 150; Gaps 22;

```

OY 14 LALLAFAAQVPTVDEL--LAMPAGEMTSTGQNDENTRHSPLQITTEVGOLOLV 71
DB 8 LAIMSIGGAQVAFANDELQKIDP--NOMAIOTGVANLRYSKLDQINKDNVSKLOVAM 66
OY 72 --ARMOGKQVQVPLHLDGVMTLNDP-----GDV----- 99
DB 67 TFSIGVLRGH--EGSPLVIGDLMVYHFEFNTVYALDLSKQGVKTEKODPNVYPMC 125
OY 100 -----IOAIDAKTGDLT-----BHRQPLNATLNVANGV 131
DB 126 CDVNRGVAAVADNKFILHQADTVVADLKATGKVLMSYKANGDAPATGNTTATVMPVKDKI 185
OY 132 IVASGTQVSPFGCFVSGHDSATGEELMRYFT-----PRAGEGDE 173
DB 186 LVGSGEGFVGRG--HYTAATSMADGKVLNMGYSMDPSDPLLDPEKTHLGRPVGKDGLT 244
OY 174 TWGNDYFARMWNG--AMQITVDPVTLNVHGSRAVAPASGTORGPGGTLYGTNTRFA 230
DB 245 TWBGD---QWKIGGGTGWWSYDDEMLVYGTGNSTMPTOR--PGDNRW-SMTIFA 298
OY 231 VRDPTGEIYMRHQTLPROMWDECTFEEMVTVNDVOSTEMEGHSINPMAATGERRYLT 290
DB 299 RQDVTGMAMKLIQMTPEHMDYDGVNEMILEQIDCK-----DKKLT 342
OY 291 GVPCKTGMQDPAETGEFL-----WARDTNQNMIE-----SIDENGI 329
DB 343 HFD--RNGGTYMDRVLTGELVAEKYDPTVNMATEVMDPSDXYGRPOVVAQSTEQNG- 400
OY 330 VTANEDALIKELDEYDVCPTFLGGRWPSALNPDSGIYFPIPLNVYCDMAVQDEFTS 389
DB 401 ---EDWT---GVCPRALGTGKQDQPAAYSPTKTELFYVPTNVCMDYEPFRVSTYA 450
OY 390 MDVNTSVNTKLPBGKDM--IGRIDAIDISTGRTLMSYERAAVNTSPVLTSGGVLENG 446
DB 451 GQPVGATLSYVTP--KDSHGMGNFAMLNKKEKIKMSLPEPFSVMSALNAGVYFYG 509
OY 447 GTDRFPALSOEGETIMOTRLATFVASGQAIETVDGMYAIAAG-----GVSGSLN 501
DB 510 TLEGYLAAVDAATGKELRKTPSGVIGVNMVYAREGQYAVLGSVGWAGIGLAAGLT 569
OY 502 S 502
DB 570 N 570

RESULT 6
ID 046444 PRELIMINARY: PRT: 708 AA.
AC 046444:
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE QUINOHAEMOPROTEIN ETHANOL DEHYDROGENASE TYPE I PRECURSOR
DE (EC 1.1.99.-) (OH-EDH1).
GN OHEDH.
OS Comamonas testosteroni (Pseudomonas testosteroni).
OC Bacteria; Proteobacteria; beta subdivision; Comamonadaceae; Comamonas.
OX NCBI_TaxId=285;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15667;
RX MEDLINE=96184549; PubMed=8654419;
RA Stoerovogel J., Kraayveld D.E., van Sluis C.A., Jongejan J.A.,
RT "Characterization of the gene encoding quinohaemoprotein ethanol
RL dehydrogenase of Comamonas testosteroni";
RN Eur. J. Biochem. 235:690-698(1996).
RP SEQUENCE OF 32-54 AND 477-490.
RC STRAIN=ATCC 15667;
RX MEDLINE=95324580; PubMed=7601151;
RA De Jong G.A.H., Geertlof A., Stoerovogel J., Jongejan J.A., De Vries S.,
Dune J.A.,

```

```

RT "Quinohaemoprotein ethanol dehydrogenase from Comamonas testosteroni.
RT Purification, characterization, and reconstitution of the apoenzyme
RT with pyroloquinoline guinone analogues."
RL Eur. J. Biochem. 230:899-905(1996).
RL [3]
RP CHARACTERIZATION.
RX MEDLINE=86242113; PubMed=3521592;
RA Groen B.W., van Kleef M.A., Dune J.A.;
RT "Quinohaemoprotein alcohol dehydrogenase apoenzyme from Pseudomonas
RL testosteroni";
RL Biochem. J. 234:611-615(1986).
RL [4]
RP CRYSTALLIZATION.
RX MEDLINE=21536088; PubMed=11679760;
RA Oubrie A., Huizinga E.G., Rozeboom H.J., Kalk K.H., de Jong G.A.H.,
RT Dune J.A., Dijkstra B.W.;
RT "Crystallization of quinohaemoprotein alcohol dehydrogenase from
RL Comamonas testosteroni: crystals with unique optical properties";
RL Acta Crystallogr. D 57:1732-1734(2001).
CC -1- FUNCTION: CATALYZES THE DYE-LINKED OXIDATION OF PRIMARY ALCOHOLS
CC TO THE CORRESPONDING ALDEHYDES AND THE (SUBSEQUENT) OXIDATION OF
CC THE ALDEHYDES TO CARBOXYLIC ACIDS.
CC -1- COFACTOR: POQ, HEME, AND CALCIUM.
CC -1- SUBUNIT: MONOMER.
CC -1- SUBCELLULAR LOCATION: PERIPLASMIC (POTENTIAL).
CC -1- INDUCTION: BY ETHANOL AND BUTANOL.
DR EMBL; X81880; CA57464.1; -.
DR HSSP; 09247; 1PQ.
DR InterPro; IPR002372; Bac_PQO_repeat.
DR InterPro; IPR003088; Cyl_C1.
DR Pfam; PF01011; Bacterial_PQO_C1C.
DR Pfam; PF00034; cytochrome_c_1.
DR PRINTS; PR00605; CYTOCHROME_C1.
DR SIGNAL; PQO; Heme; Calcium; Oxidoreductase; Periplasmic.
KW CHAIN 1 31
FT 32 708
FT BINDING 635 635
FT FT 638 638
FT METAL 639 639
SQ SEQUENCE 708 AA; 76822 MW; 99AB54BD66ACAB3 CRC64;

Query Match 17.8%; Score 479.5; DB 2; Length 708;
Best Local Similarity 25.0%; Pred. No. 9.3e-27;
Matches 148; Conservative 84; Mismatches 238; Indels 121; Gaps 18;

OY 3 PTLNASSAGALALAAPFAQVPTV-----DELANPP-AGEMISTYQNDEN 50
DB 12 PGRMVLWLAAGL--SAAFNQTPAQAQAAVAVRVGDFTFRAVARTPMPPIGVDAE 69
OY 51 YRHSPLQITTEVGOLOLVNARKMQPK-VQVPTLHLDGVMTLNDPQIDAKTG 109
DB 70 TRSRLDQINAAVADLGLAWSYMLESTRGVLEAPVVDGIMVVSASVVAHIDRTGN 129
OY 110 LWEHRQDLPNT--ATLVANG----- 130
DB 130 RIWYDPOIDSTGFGKGCDDVNNGVALLMKGYVCAAMDRLIALDAATGKEWHQNTFE 189
OY 131 -----VVASSTCOYSPFGCFVSGHDSATGEELMRYFT----- 165
DB 190 GOKSLITTGAPRVFKKVTIIGKGALEYGRG-YIAYDAETGKKMWFVPGDPSPKPF 248
OY 166 --RAGEGDETWGNDYFARMWNGA-----NGQITVDPVTLNVHGSRAVAPASGTORG 218
DB 249 EDESMKRAARTW--DPSGKMWEGGGCTWMDSMITDAELNLTGCTGSGFWSHKVRS 306
OY 219 GGLTGTNTRFAVBDPTGELVNRHQTLPROMWDECTFEEMVTVNDVOSTEMEGHSIN 278
DB 307 GGDNIYLAIVALLDPTGKYKWHYQETPGDMDYTSQPTMLADIKI----- 353

```

[illegible]

RESULT	7		
09F9U2		PRELIMINARY:	PRT: 695 AA.
ID	09F9U2;		
AC	01-MAR-2001 (TREMBLrel. 16,	Created)	
DT	01-MAR-2001 (TREMBLrel. 16,	Last sequence update)	
DT	01-JUN-2001 (TREMBLrel. 17,	Last annotation update)	
DE	ALCOHOL DEHYDROGENASE.		
GN	ADH.		
OS	Pseudomonas stutzeri (Pseudomonas perfectomarina).		
OC	Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;		
OC	Pseudomonas.		
OX	NCBI_TaxID=316;		
RN	(1)		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=BC-2; Herrick J.B., Okinaka R.T., Brainard J.B.,		
RA	Chang C.-H.,		
RA	Pervilliger T.C.;		
RT	"Identification and characterization of genes activated by 2-		
RT	chloroethanol in Pseudomonas stutzeri BC-2.";		
RL	Submitted (Aug-1999) to the EMBL/GenBank/DBJ databases.		
DR	EMBL: AF176640; AAC09249.1; -.		
DR	HSSP: Q9Z4J7; 1FLG.		
DR	InterPro: IPR002372; Bac_PQQ_repeat.		
DR	InterPro: IPR000345; CylC_heme_bind.		
DR	InterPro: IPR003088; CylC.		
DR	Pfam: PF00101; Bacterial_PQQ: 6.		
DR	Pfam: PF00034; Cytochrome_C: 1.		
DR	PROSITE: PS00190; CYTOCHROME_C: UNKNOWN_1.		
DR	PROSITE: PS00190; CYTOCHROME_C: D29A698A642E1A83 CRC64;		
DR	SEQUENCE 695 AA; 75842 NM;		

```

Query Match      17.8%: Score 477.5; DB 2: Length 695;
Best Local Similarity 26.2%: Pred. No. 1,3e-26; Indels 111; Gaps 19;
Matches 147; Conservative
OY      15 ALLAPAPAOYPTVTDLL--ANPPAGEMISYGOONENRHSPLTQITTEYGOLOLYMA 72
      16 ALIYA-AGAQAAXDEALAIASBODSEWLSHGRIYAQEFSPSLKOIDAGNNGKGLAMY 74
OY      73 RGMOPCK-VQVTPRIIDHGVNVLAMPDVIQALIDGDLINENRQLP----- 119
      75 LDLENRRGLEATPLVSDGVYLAISLSRWAAVDLRSGRLMQDPDYDRSHSRYTCODAP 134
      120 --NIATLN-----IYANGVITAG 135
OY      135 NRGVALMGKVVYVYGAIDGRLIALDACTGRELMEQTTDPKAPYSITGAPRVYAGKYLIGN 194
      136 STCOQTSFGQGFSGHDSATGELMENNFIIPRAC-----EGDETW-GNDYEAKW 183
OY      195 GGAELGAVRG-FFSAIYDAETGKAMAFYVVP--GPPAPYIEHPLMAEAKTKWQDOY--W 248
      184 MFG-----AMGQITDYPVNLNHHGSAVGAAPASEYQRTPGGTLTGYNIRFANRBDTGEIV 239
OY      249 KLGGGGYWMSDAMAYDPELDILYTGIGNSPMMRREIRSPGGGDMILYSSILALRDPGKLL 308
      Db      249 KLGGGGYWMSDAMAYDPELDILYTGIGNSPMMRREIRSPGGGDMILYSSILALRDPGKLL 308

```

[illegible]

RESULT	8		
09AO48		PRELIMINARY:	PRT; 629 AA.
ID	09AO48;		
AC	09AO48;		
DT	01-JUN-2001 (TrEMBLrel. 17,		Created)
DT	01-JUN-2001 (TrEMBLrel. 17,		Last sequence update)
DT	01-DEC-2001 (TrEMBLrel. 19,		Last annotation update)
DE	METHANOL DEHYDROGENASE MXAF.		
GN	MXAF.		
OS	Methyllobacterium nodulans.		
OC	Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;		
OC	Methyllobacterium group; Methyllobacterium.		
NCBI_TaxID=114616;			
OX	[1]		
RN	SEQUENCE FROM N.A.		
RP	STRAIN-OR52060;		
RC	MDLINL-20566686; PubMed-11114919;		
RX	Sy A., Giraud E., Jourdan P., Garcia N., Williams A., de LaJudie P.,		
RA	Pin Y., Neyra M., Gillis M., Boivin-Masson C., Dreyfus B.;		
RA	"Methyloleptophilic Methyllobacterium bacteria nodulate and fix nitrogen in		
RT	symbiosis with legumes."		
RL	J. Bacteriol. 183:214-220(2001).		
HL	EMBL: AF20764; AAC49450.1; -.		
DR	HSSP: P38539; AAAH.		
DR	InterPro: IPR001475; Bac_PQQ.		
DR	InterPro: IPR003732; Bac_PQQ_repeat.		
DR	pfam: PF01011; Bacterial_PQQ_7.		
DR	PROSITE: PS00363; BACTERIAL_PQQ_2: 1.		
DR	PROSITE: PS00364; BACTERIAL_PQQ_2: 1.		
DR	SEQUENCE 629 AA; 68798 MM; 3DB55DB373BBOB2 CRC64;		

	Query Match	17.5%	Score 470.5:	DB 2:	Length 629:	
	Best Local Similarity	25.1%:	Pred. No. 3.6e+26:	Indels 139:	Gaps 22:	
	Matches 150:	Conservative 97:	Mismatches 211:			
OY	4 TSLMASAGALLAAAPFAOVPTVDEL-ANPAGEWISGONENRHSPLQIYTE 66					
	I : I :	I : I :	I : I :	I : I :	I : I :	I : I :
Dd	11 TGVSVAALALPLPGS-----ALMDKILVELSKSDGWPMKANDSNYSKLQINAE 99					
OY	63 NVCOLDLYW--ARGMOPEGVOVTFPIIDHGYM-----LANPQDY----- 125					
	I : I :	I : I :	I : I :	I : I :	I : I :	I : I :
Dd	67 NKKMLKWSOFSTGLINH--EGAPLVVDIGIMYYHTSFNNITFALGDLPDKILMODKPKQ 121					
OY	100 -----IQATDAKTGDLIWEHRQLPIN 185					
	I : I :	I : I :	I : I :	I : I :	I : I :	I : I :
Dd	126 NPAAASVACDLYNRGLAWPEDGKITPSLILKTLTLDGHVALINAADTGTEVTKINDSINV 172					
OY	122 -TTUNI-----VANGVIYASTCQYSPBECFYSGSDSATGEELIRNFYIPRAGEBCD----- 241					
	I : I :	I : I :	I : I :	I : I :	I : I :	I : I :
Dd	186 GSTLIIVAYVKDKVILIGSSGAELVRG-YLFADVAVDTGQWKRAY---ATGPSDDLILA 241					

QY 173 -----ETWGNDEYKRMGTG---WGOITVDPVNLVHGSTAVGPA52 212
 Db 242 KQENTHNAHYGOKGLSTWEGD---AMKIGGGTMMGWYADPGTINLITFGTGPAPWNE 298
 QY 213 TQRPGGTTLGTRPAPVDPDTGELVNRHQTLPDRNDQECTEFMAVTVNDVOPSTEME 272
 Db 299 TMR---PDKNW---TWTFIARDVDTEKFKGYOKTPHDEMDYAGVNMML-----STOKD 348
 QY 273 GLOSINNAATGERRVLTGVPCKTGMOPDAETGEFLMAR---DNYNMIESIDENGI 329
 Db 349 -----RSGKERKLLTHPDRNGIYTTLDRTNGDIDISAKIDT---VNFERYDLKSG 397
 QY 330 VVNEDEAILKEID-VEYVYCEPTFGGNDPSPALNPDGSIYFPLNNVCYDMAVDOEFTS 388
 Db 398 LPRVDEYGTBRMDHLAKICPSAMGYHNOCHSDYDEPRLTFYMGINHCMDWEPFLPYR 457
 QY 389 SMQVYTSNVTKLPPGK-----DMIGRIDAIDISTGRTLSVERAANYSPLYSTGGV 442
 Db 458 AGQFVGAATLMNTPGPKGDBQNAEGQIKAYDALITGFKWKEKERRAVMGGLTATAGNV 517
 QY 443 LFNQSTDRYFRALSQETGETLMOTRLATVASGAISTEVDGMQYVAAGSGVYSG 499
 Db 518 VFTGLDGFIAKRSHTGELLMKALPSGALGYVPTVTHKGTQYVAT-----YGVG 569
 RESULT 9
 P71509 PRELIMINARY: PRT: 601 AA.
 AC P71509;
 DT 01-FEB-1997 (TREMBLrel. 02, Created)
 DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)
 DE METHANOL DEHYDROGENASE LARGE SUBUNIT HOMOLOG.
 GN MKAP.
 OS Methylobacterium extorquens.
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OX Methylobacterium group; Methylobacterium.
 RN NCBI_TaxID=408;
 RP SEQUENCE FROM N.A.
 RC STRAIN=AM1.
 RX MEDLINE=97312011; PubMed=9168622;
 RT "Cholesterol L., Lidstrom M.E.;
 RT "Molecular and mutational analysis of a DNA region separating two
 RL methylobiotrophy gene clusters in Methylobacterium extorquens AM1.";
 DR EMBL: U72662; AAB58890.1;
 DR HSP: P38539; AAAH.
 DR InterPro: IPR002372; Bac_PQQ_repeat.
 DR Pfam: PF01011; Bacterial_PQQ: 7.
 SO SEQUENCE 601 AA; 64952 MW; 68E45C7059CBB239 CRC64;

Query Match 17.2%; Score 463.5; DB 2; Length 601;
 Best Local Similarity 25.4%; Pred. No. 1.e-25;
 Matches 152; Conservative 97; Mismatches 220; Indels 129; Gaps 22;

QY 1 MKPTSLMASAGALALLAPARAQVTPYTDLLANPAGEMISXGNOENYRHSPLQIT 60
 Db 1 MRAVHLALGAGIAA--ASPALANESYLKQ--YANP--AAQVLTQVYANTRYSLQIDN 54
 QY 61 TEWVGQLOLW--ANGMOPKQVQVTPPLIHGVNLANP----- 96
 Db 55 ASNVKRLQVAMTSTGVLRGH-EGSPLVGNIMVTTPPPNIVVALDLDQAKIVMKYEP 113
 QY 97 -----GDVIOALDAKTGDLIMHRQLP-----N 120
 Db 114 KQDPSVIFVMCDDTVNRGLAVADGAILLHQADTTLVSLDAKSGKNNYSVKNGBDSGKFN 173
 QY 121 INTLNIANGVIVAGTCQYSPFGCFVSGHDSATGEELMRNYF----- 164
 Db 174 TATVLPKDKVIVYISGSGEG--VOCHVTAAYDLKSGKVMKGTISGDDQDLIVPEKTEL 232

QY 165 --PRAGEEGDETMGNDYKRMGTG---ANGQITVDPVNLVHGSTAVGPA52 219
 Db 233 GRPIAKDSLSLTWEGD---OMKTGGCGTWGMSPTPKDLMTYSG--NSTYNPKQRBG 287
 QY 220 GLVGTNTRAVREDGEIYWRHQTLPDRNDQECTEFMAVTVNDVOPSTEMEGLQSI 279
 Db 288 DNK--SMITWARRPDDGMKAWYQMTPHDEMDYAGVNMML-----OKRDG----- 334
 QY 280 NAATGERRVLTGVPCKTGMOPDAETGEFLMARDT-YONMIESID-ENGIVYNEAI 337
 Db 335 -----KDRPLTTHD-RNGRGYITLDPATGEVLYAEKEPVPVNNATKVDLDKSGITGRPLV 389
 QY 338 LKELDVEYD-----VCPFTFGGRDPSPALNPDGSIYFPLNNVCYDMAVDOEFTS 389
 Db 390 VSKYSTEDQGEVLSKGTICPALCTKQCPAPSPKGTGTFVPTVNHVCMDEPPRYTTP 449
 QY 390 MDVYTSNVTKLTP-PGR-DMIGRIDAIDISTGRTLSVERAANYSPLYSTGGVLENG 447
 Db 450 GQPIVGAATLMNTPGPKGDBQNAEGQIKAYDALITGFKWKEKERRAVMGGLTATAGNV 517
 QY 448 TDRYFRALSQETGETLMOTRLATVASGAISTEVDGMQYVAAGSGVYSG 500
 Db 510 LEGFLKAVDSKTGKELKRTKTSGLIGNVMTYEHKGRQHAVLSTGVGNAGIGLAAGL 567

RESULT 10
 O9AP95 PRELIMINARY: PRT: 691 AA.
 AC O9AP95;
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DE 1-BUTANOL DEHYDROGENASE BDH.
 OS Pseudomonas butanovora.
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 OX Pseudomonas.
 RN NCBI_TaxID=86174;
 RP SEQUENCE FROM N.A.
 RC Vangnai A.S., Arp D.J., Sayavedra-Soto L.A.;
 RT "Characterization of the expression of two distinct alcohol
 RT butanovora.";
 RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF355798; AAK27220.2;
 SO SEQUENCE 691 AA; 75070 MW; 4FC7ED20CDA14E64 CRC64;

Query Match 17.2%; Score 461.5; DB 2; Length 691;
 Best Local Similarity 26.4%; Pred. No. 1.9e-25;
 Matches 145; Conservative 82; Mismatches 205; Indels 117; Gaps 19;

QY 39 GEMWISYGMQENYRHSPLQITTEENYGOQLVWARGMOPK-QVTPPLIHGVNLANP 97
 Db 41 GEMRHGIDYDAGTRYSPLAQITPDNAKELGLWVSYLESSRGVEARPIYVDGVMYTAW 100
 QY 98 DVIOALDAKTGDLIMHRQLP-----NATLNIANGVIV-ASTQYSPFGCFVSGHDS 152
 Db 101 SVVHALDVRSGRKLMYVDEYPEREKKNACCDYVNRGVANHEKRYVGSGLDRIVA-IDA 159
 QY 153 ATGEEELM-RNYF-----PRA--GEBEDETWGNDEYKRMGTGAM-----GQ 190
 Db 160 RTGRKWRERNLIDDDKPYTTGAPRVIRKRVYVINGGAKEFCVGRYITAYDPTAASRPV 219
 QY 191 I-----TYDPTVNLVHGSTAVG----- 208
 Db 220 VPGPDDPSLPEDASMAAKATWDPAGVYLGSGRRRHGVELDGLYKAGCCCSABATPS 279
 QY 209 PASFQRTGPGTILYNTFAVVRPDTGEIYWRHQTLPDRNDQECTEFMAVTVNDVOP 266
 Db 280 PWSHKKSPAGGDMLYTASIVALLRPDTGEIYWRHQTLPDRNDQECTEFMAVTVNDVOP 266
 QY 267 PSTMEGLQSIINNAATGERRVLTGVPCKTGMOPDAETGEFLMARDTNYONMIESIDE 326

Fri May 24 11:27:28 2002

walick-934-125.pcp.rsp

page 7

```

Db 340 P-----RKVILHAP-KNGFEFFIDRDKFKTSQNPVPMATGDE 380
OY 327 NGIVYVN-EDALIKELDYEVCEPFLGGRDPSAALNPDSTGYIFPLNVCYDKMAAVDO 385
OY 381 NGRPIENPGAMPGLHSMRF---PAPSARTIMHMSYSPTQGLAYEPANITPL-VLOEDK 436
OY 386 EFTSMQVNTSN-----VTKLPKGMKIDIGIDAIDISTGRTLMSVERAAA 430
Db 437 NWS-----YNOAQPGQAMAGIGMNLGMLVNPBPASQPGRLIANDPVQKEVNRKEHVSF 492
OY 431 NISPVSTGCVLENGGDRYFRALSOETGETLMOTRLATYASQAISYEDVMQYVIA 490
Db 493 WNGTLYTLAGNVFGSTADARILALDARDGKELMSAPMGTVIAPVAYEVDGKQYVIA 552
OY 491 -GGGVSYGS 498
Db 553 VGMGVYGN 561

```

```

RESULT 11
ID 024759 PRELIMINARY; PRT: 633 AA.
AC 024759;
DT 01-JAN-1998 (Tremblrel, 05, Created)
DT 01-JAN-1998 (Tremblrel, 05, Last sequence update)
DT 01-DEC-2001 (Tremblrel, 19, Last annotation update)
DE METANOL DEHYDROGENASE ALPHA SUBUNIT PRECURSOR (EC 1.1.99.8).
GN MXAF.
OS Hypnomicrobium methylotrochum.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Hypnomicrobium group; Hypnomicrobium.
OX NCBI_Taxid=84;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-GM2;
RC MEDLINE-97457202; PubMed-9311140;
RX Tanaka Y., Yoshida T., Matanabe K., Izumi Y., Mitsunaga T.;
RA "Cloning and analysis of methanol oxidation genes in the methylotroph
RT Hypnomicrobium methylotrochum GM2."
RL FEMS Microbiol. Lett. 154:397-401(1997).
DR EMBL; AB004097; BAA23272.1; -.
DR HSSP; P38539; 4AAH.
DR InterPro; IPR002372; Bac_POO.
DR Pfam; PF01011; Bacterial_POO_7.
DR PROSITE; PS00363; BACTERIAL_POO_1; 1.
DR PROSITE; PS00364; BACTERIAL_POO_2; 1.
DR Signal; Oxidoreductase.
KW SIGNAL.
FT CHAIN 35 633 POTENTIAL.
FT SEQUENCE 633 AA; 69853 MW; B47A23A2779E2C1B CRC64;
SQ

```

```

Query Match 16.0%; Score 431.5; DB 2; Length 633;
Best Local Similarity 24.1%; Pred. No. 2.6e-23;
Matches 146; Conservative 96; Mismatches 214; Indels 151; Gaps 22;

```

```

OY 6 LMASAGALLALAAPFAOYVPTDELLANPAGE-WISYQONENYRHSPLQTITENY 64
Db 12 LMSASCAIYALOVGVASSAYANLKLLELSKSNEMWMPKKNDSNYSSTQVNAENY 71
OY 65 GOLDIYAROMQGVK---QVPLIHIDVYVY-----LANGDV----- 99
Db 72 KOLHAMS--FSTGELHGHGAPLYGDMVYHSSFPNKTALINDGHLIMHSPKOD 129
OY 100 -----IQADAKTGDLIWEHRQLEPNIA 122
Db 130 PAASVACCDLVNGLAYWPGDKTPALYKIQDDHVALNAKKGEEFMK----- 180
OY 123 TLTNIYANGVYAGTCQYSPGCE-----VGHGSA-----TGEELMNYET- 164
Db 181 ----VENQDIRVGTITQAPYVHDLAIVGSSGAEIGVRGHVATYAVKKGEDAMWYATG 236

```

```

OY 165 -----PRAGEE--GDETWNDEYARWMEGA---WQITYPDTNINLVHGS 204
Db 237 PDEILADDFNSANPHYGOKGLGTATWEGD---AMKIGGTNMGWATYDPAQANLIYGS 293
OY 205 TAVGASEFQGRPGGTLYGTNTFRAPVBDTGEIYWRHQTLPENNMDECTFEKMYATND 264
Db 294 GNPAPMETWR--PGDNKK--TMTITARADDTGKMKFGYOKTPHEMDFAGVNYTMSLE-- 348
OY 265 VQSTEMEGLOSINPNAATGERAVLTGVCPCKTGTMQFQDAETGEFLMA--RDTNYONMI 321
Db 349 ---QDKE-----GKKRKLTHPDRNGIYTTLDRENGDLISADKIDDD--VNVF 392
OY 322 ESDINGIVYNEDALIKELDEVE-YDVEPFLGGRDPSAALNPDSTGYIFPLNVCYDM 380
Db 393 KHVDSKGLPYRDEPFGTRMDHKGTEICPSAMGYHNGHDSYDPTKOLFPGINNICMDW 452
OY 381 MAVDOFTSMQVNTSNVTKLP--PGKDM-----IGRIDAIDISTGRTLMSVERAANSP 434
Db 453 EEFMLPYRAGQFVFGATLMTYPGKGRONYLIGQIKAYNAITNTYKKEHMERFSWVG 512
OY 435 VLTSGGVLENGGDRYFRALSOETGETLMOTRLATYASQAISYEDVMQYVIAAG-G 493
Db 513 TLTATGNLVYFTITDGLFKARNSDTGELIMRKLDPGCVIGTPYTHEKGYQYIAVMSYG 572
OY 494 VSYGSL 500
Db 573 GMPGVGL 579

```

```

RESULT 12
ID 09AGW3 PRELIMINARY; PRT: 623 AA.
AC 09AGW3;
DT 01-JUN-2001 (Tremblrel, 17, Created)
DT 01-JUN-2001 (Tremblrel, 17, Last sequence update)
DT 01-DEC-2001 (Tremblrel, 19, Last annotation update)
DE 1-BUTANOL DEHYDROGENASE BOH PRECURSOR.
OS Pseudomonas butanovora.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_Taxid=86174;
RN [1]
RP SEQUENCE FROM N.A.
RA Sayavedra-Soto L.A., Vangnai A.S., Arp D.J.;
RA "Characterization of the expression of two distinct alcohol
RT dehydrogenases involved in butane metabolism in Pseudomonas
RT butanovora."
RL Submitted (DPC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF326086; AAK15506.1; -.
DR HSSP; Q924J7; 1FIG.
DR InterPro; IPR002372; Bac_POO_repeat.
DR Pfam; PF01011; Bacterial_POO_4.
KW SIGNAL.
FT CHAIN 1 29 POTENTIAL.
FT SEQUENCE 623 AA; 67553 MW; A573A1C64AC12D55 CRC64;
SQ

```

```

Query Match 15.9%; Score 428; DB 2; Length 623;
Best Local Similarity 24.9%; Pred. No. 4.6e-23;
Matches 157; Conservative 81; Mismatches 217; Indels 176; Gaps 24;

```

```

OY 2 KPTSLMASAGALLALAAPFAOYVPTDELLAN--PRAGENISYQONENYRHSPLQI 59
Db 7 KPFALRAIVVATAALSLIPA-AAVTDTWEDINDHDKRTGDTVLTGLGKAORHSPFLKAI 65
OY 60 TTEVNGOGLIWA-----RGMQPKYOVPTLIHDGWTYLANPDPYIOAIDAKTGDLI 111
Db 66 NTDNVAANLVPAKSPSPGEGKORGE---AOV--LVHDGVYATASYSRIPALDARGKRL 120
OY 112 WEHRQLEPN--TATLNIYANGVYAGTCQYSPGCEVSGHSDATGEELMNYF----- 163
Db 121 WETNARLPDIRPCDDVYVNGAATYIGDKVFTGTDAMVALLDRKTGKVVYVKKRFGDHKG 180

```


Query Match	15.6%	Score 419;	DB 2;	Length 599;
Best Local Similarity	24.4%	Pred. No. 2e-22;		
Matches 142;	Conservative 96;	Mismatches 227;	Indels 118;	Gaps 21
QY	4	TSLLMSAALMLLPAFAQVPTVDLLNPPAGEWISYGVQNGENRRSPLOTTER	63	
Db	6	TATGTFVADLIALSMLPSVPAASDLSEILGALNP--NNMTQTGDGYQGHLSKLSQTLTTON	63	
QY	64	YQQLD--VAARGMQPKQVQVTPLLIHGVVYLANP--PGVIGALDAKTKGDLIWEHR--	115	

```

Db      64  YKNLESGHGHSSTGLLGH--EGAPLYIGDMATINPEPNNFALMLADEPKITWQHPKQ 122
OY      116 -RQLENTATLTIYANGV-----IVA-----GSTCY 140
Db      123  DASVAAVACCDIYNGRLAYDDGHIIFKTQDGLHLVASDAKTGKELMKMENDCPAVGSTITQ 182
OY      141  SPF-----GC-----FVSHSDANTEELMRNYFI-----PRA 167
Db      183  APEAVKGLVLYGSGALFGRVITAYADQKGTGELMWSFAGDPNEPLNAKFNKONPH 242
OY      168  GEE--GDEITGNDY EAKRMGTGA--WQIITDPVNLVHYSTAVGASPTSGTGGTL 222
Db      243  GQGLGKTKTWEEG--GMKIGGGTNMGWYAVDPKLNLFYSSGNAPAMNMTM--PGDK 297
OY      223  YGNTFFAARPDPTGETIWMHOFLEPDMNDQCETEMMATVNDVOPSTENEGLOSIINPA 282
Db      298  W-TMTIWARDVDTGAANKXGYKTPHDEHDPAGVNOMLTIOAANGTKO----- 344
OY      283  TGBRRLLYGPCKTITMOPALETGEFLMARDIN-YONMIESIDENGIYVEMDILAEI 341
Db      345  -----PLTHVD--RNSINITLNRQTSITVOAAKVDPAVNNFFKKVLDLKTGLPVADPEFSRM 399
OY      342  DVE-YDVCPTEFLGGRDMPALINDGSITYPEIPLANNCYDMAVAODEFTSMDYNTSNYTK 400
Db      400  DHRGTNIPDSAMGFHNGDLAYDPDSTYFTGLNHCCKMDEPFMLPYRAGOFFGVCATIAM 459
OY      401  LP-----PCKDMIGLIDAIDISTGRTLSVYERRAANYSPVLSGGGVLNPGSGTDYPAFLS 456
Db      460  YPGNGPTKKMGVGLAMDQSTGEVWITWMEKFSYWGVLATKGLVLEYNNLDDNITALL 519
OY      457  QETSETLMQRLATVASGAISIEYVDMGYVAILAGGVSYSG 499
Db      520  KTNCKEIMFKMGSAGIAGPMAYVAKRGKYIA-----TINVG 557

```

```

RESULT 14
0934G0
ID      0934G0
AC      0934G0;
PR      PRELIMINARY;
PT      PRT;
SA      695 AA.
DT      01-DEC-2001 (TREMblrel. 19, Created)
DT      01-DEC-2001 (TREMblrel. 19, Last sequence update)
DE      01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE      LUPANINE HYDROXYLASE PRECURSOR.
OS      LUPH.
OC      Pseudomonas sp. DH2001.
OC      Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OX      Pseudomonas.
NX      NCBI_TaxID=163360;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      Hopper D.J., Kaderbhai M.A., Little A.R., Marriott S.A., Young M.,
RA      Roczinski J.;
RT      "Cloning, sequencing and analysis of the gene for lupanine
RT      hydroxylase, a quinocytochrome c from a Pseudomonas sp.";
RL      Submitted (JUN-2001) to the EMBL/Genbank/DBD databases.
DR      EMBL; AJ318095; CAC6440.1; .
KW      Signal.
FT      SIGNAL.
FT      CHAIN.
FT      1
FT      26
FT      27
FT      695
FT      695
FT      LUPANINE HYDROXYLASE.
FT      SEQUENCE 695 AA; 74956 MW; 0996C6387E457FE0 CRC64;

```

Query Match	15.08;	Score 404.5;	DB 2:	Length 695;
Best Local Similarity	25.38;	Pred. No. 2,8e-21;		
Matches 136;	Conservative 73;	Mismatches 216;	Indels 113;	Gaps 18;
QY	38 AGENISYGQNEQNRHSPPLQTITTEVVGQLDVMARQGPKRVGV	-TPRLHDGMYLNP	96	
Dd	36 SGMSNLIGGGNEGHEYSALDKDYNSVKNNKLGSLWPTDEADGDLYGNPLVDADVIVOGGP	95		
QY	97 GDVIOALDATGTGLLIHRRL-----PNIATLINIANGCVIAGSTCG-----	139		
Dd	96 PGRIVANDLKRTGKNLTFTYPEVOYDKDTSMTGFNGFTHVNGLAIVDDNDNYISSGYCTLLAV	155		

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OW protein - protein search, using sw model

Run on: May 24, 2002, 10:19:18 : Search time 101.42 Seconds
(Without alignments)
553.069 Million cell updates/sec

Title: WALICK-934-128.PEP

Perfect score: 2705

Sequence: 1 MKPTSLIMASAGALALALP.....GMQVYALAGSVSYGSGINS 505

Scoring table: BIOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_032802:*

1: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT:*

2: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT:*

3: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT:*

4: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT:*

5: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT:*

6: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT:*

7: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT:*

8: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT:*

9: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT:*

10: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT:*

11: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT:*

12: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT:*

13: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT:*

14: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT:*

15: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT:*

16: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT:*

17: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT:*

18: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT:*

19: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT:*

20: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT:*

21: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT:*

22: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2669.5	98.7	579	19	AAW37873
2	2344.5	86.7	579	19	AAW37874
3	2310	85.4	579	19	AAW37876
4	2251	83.2	578	19	AAW37875
5	480.5	17.8	754	21	AAW35987
6	479	17.7	738	13	AAW20192
7	478	17.7	738	13	AAW13993
8	476.5	17.6	742	11	AAW05235
9	314	11.6	740	20	AAW5019
10	155.5	5.7	443	22	ABG24430
11	134.5	5.0	824	18	AAW09614

12	134	5.0	715	22	AAW81738
13	133.5	4.9	910	22	AAW83007
14	133	4.9	696	22	AAW82914
15	133	4.9	1016	22	AAW34349
16	130	4.8	470	22	AAW82701
17	129	4.8	948	22	ABG25909
18	129	4.8	1510	22	ABG21573
19	117	4.3	593	11	AAW07014
20	116	4.3	480	8	AAW70468
21	115.5	4.3	594	12	AAW10005
22	113	4.2	404	22	AAW82772
23	110	4.1	269	11	AAW07006
24	110	4.1	269	15	AAW53295
25	109	4.0	269	10	AAW94788
26	109	4.0	593	15	AAW62944
27	108.5	4.0	530	22	AAW12236
28	107.5	4.0	886	22	AAW39966
29	107	4.0	886	21	AAW10645
30	106.5	3.9	2599	21	AAW75098
31	106	3.9	501	18	AAW10644
32	106	3.9	514	22	AAW95325
33	106	3.9	770	22	AAW30824
34	106	3.9	966	19	AAW49077
35	105.5	3.9	429	22	AAW66542
36	105.5	3.9	469	21	AAW56578
37	105.5	3.9	429	22	AAW3799
38	105	3.9	501	18	AAW82625
39	105	3.9	501	18	AAW26108
40	105	3.9	501	18	AAW26109
41	104	3.8	501	18	AAW26107
42	103.5	3.8	493	22	AAW33026
43	103.5	3.8	649	22	AAW66538
44	103.5	3.8	974	19	AAW49079
45	103.5	3.8	1376	22	ABW52592

ALIGNMENTS

AAW37873	1	AAW37873 standard; Protein; 579 AA.
AC	AAW37873:	
XX	10-AUG-1998	(first entry)
DE	Alcohol and/or aldehyde dehydrogenase A amino acid sequence.	
XX	Alcohol/aldehyde dehydrogenase A enzyme; recombinant organism;	
KW	aldehyde; ketone; carboxylic acid; L-sorbose; D-sorbitol;	
KW	2-keto-L-gulonate acid; L-ascorbic; inhibition.	
OS	Glucanobacter oxydans.	
XX	Key	Location/Qualifiers
FT	Peptide	1..23
FT	Protein	/note="signal peptide"
FT		24..579
FT		/note="mature protein"
PN	EP832974-A2.	
PD	01-APR-1998.	
PF	11-SEP-1997:	97EP-0115801.
PR	19-SEP-1996:	96EP-0115001.
PA	(HOFF) HOFFMANN LA ROCHE & CO AG F.	
XX	Asakura A, Hoshino T, Ojima S, Shinjoh M, Tomiyama N;	

S. epidermidis ope
S. epidermidis ope
S. epidermidis ope
Staphylococcus aur
S. epidermidis ope
Novel human diagno
Novel human diagno
Protein G variant
Sequence of polype
Streptococcus GX78
S. epidermidis ope
Protein G variant.
IgG-binding Strept
Protein G variant.
Streptococcus Prot
Human PRO340 poly
Human protein sequ
Human protein L1 pro
HPV6 mutant L1 pr
Neisseria meningit
Human protein sequ
Amino acid sequenc
Solium tuberosum
Human interferon-a
Human protein sequ
Human prostate can
S. epidermidis ope
HPV11 mutant L1 pr
HPV11 mutant L1 pr
C glutamicum prote
Drosophila melanog
Solium tuberosum
Escherichia coli p

DR WPI: 1998-195228/18.
 XX N-PSDB: AAV29051.
 PT Recombinant Gluconobacter oxydans alcohol and/or aldehyde
 PT dehydrogenase enzyme(s) - useful for converting substrate(s), e.g.
 XX L-sorbose or D-sorbitol to 2-keto-L-gulononic acid
 PS Claim 1; Pages 35-37; 59pp; English.
 XX
 CC This is the amino acid sequence for the Gluconobacter oxydans
 CC alcohol and/or aldehyde dehydrogenase A' enzyme. The enzymes
 CC or recombinant organisms can be used to convert suitable substrates
 CC to aldehydes, ketones or carboxylic acids, especially to convert
 CC L-sorbose or D-sorbitol to 2-keto-L-gulononic acid, which can be
 CC converted to L-ascorbic acid by standard procedures. The derivatives
 CC of ADH enzymes have desired substrate specificity, higher affinity
 CC to a substrate, lower affinity to an inhibitory compound, higher
 CC stability against temperature and/or pH and higher catalytic speed.
 XX
 SQ Sequence 579 AA;

Query Match 98.7%; Score 2669.5; DB 19; Length 579;
 Best Local Similarity 90.8%; Pred. No. 2.3e-227;
 Matches 505; Conservative 0; Mismatches 0; Indels 51; Gaps 1;

OY 1 MKPTSLIMASAGALALAAPAQAQVPTDDELANPAGEMISYQONQENRHSPLTQIT 60
 DB 1 mkptslimasagalalAAPAQAQVPTDDELANPAGEMISYQONQENRHSPLTQIT 60
 OY 61 TENVGQLOLVMAKMGKQPKVQVTPPLIHGQVWYLANPGVQIADAKTGLIWEHRQLPN 120
 DB 61 tenvgqlqlvwarqmgpkvqvtpplihdgvywylanpgvdiqaidaktgdlwehrrqlpn 120
 OY 121 IATLNSFG-----I 129
 DB 121 iatlnsfgeptrmalylgtvnyfsvwdhnlvaldtatgvtfdvdrqggedmvsnsqpi 180
 OY 130 VANGYVAGSTCOYSPFCEVSGHDSATGEBELRNRFIPRAGEGDEMGNYEARMTG 189
 DB 130 vangyivagstcoyspfcevsghdsatgelnrnfipragegdegmngnyearmtg 189
 OY 181 vanyivagstcoyspfcevsghdsatgelnrnfipragegdegmngnyearmtg 240
 OY 190 AMGOITPDTNLVHGSTAVGASSETQGTGEGTLYGNTREAVRPDTGETVWRHQTLP 249
 DB 241 amgiltvdpdtnlvhgstavgsasetqgtggtgelygntrfavrpdtgetvwrhqlp 300
 OY 250 RDNMDQCTEEMWNTVNDVQPTSTEMGLQSIINPNAATGERRVLGVPCKTGTMOFDAET 309
 DB 301 rdnmdqctefemwntvndvqptstemeglsinpnatgerrvlgvpcktgtmofdaet 360
 OY 310 GEFLLARDNYONMIESIDENGIVVNEBALIKELDVEYDVCPTFLGGRDMPASALNPS 369
 DB 361 gefllardnynmiesidengivvnedalikeldveydvcptflggrdmpasalnps 420
 OY 370 GIYFIFLANNVCYDMAVDOETSMQVNTSNVTKLPPGKMGIGITDIDISTGRITMSVE 429
 DB 421 giyfiflannvcydmavdoetsmqvntsnvtklppgkmgigitdidistgrltmsve 480
 OY 430 RAAANVSPISTGGGVLNCGSDXFRFALSQETGETIMQTRLATVASGQAISEVVGMOY 489
 DB 481 raaanvspistgggvlncgsdxfrfalsqetgetimqtrlatvasgqaisevvgmoy 540
 OY 490 VAIAGGVSYSGSLNS 505
 DB 541 valaggyvsygslns 556

RESULT 2
 AAM37874
 ID AAM37874 standard; Protein; 579 AA.
 XX
 AC AAM37874;
 XX

DT 10-AUG-1998 (first entry)
 XX
 DE Alcohol and/or aldehyde dehydrogenase A' amino acid sequence.
 KW
 KW Alcohol/aldehyde dehydrogenase A' enzyme; recombinant organism;
 KW aldehyde; ketone; carboxylic acid; L-sorbose; D-sorbitol;
 KW 2-keto-L-gulononic acid; L-ascorbic; inhibition.
 XX
 OS Gluconobacter oxydans.
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..23
 FT Protein /note- "signal peptide"
 FT Protein 24..579
 XX /note- "mature protein"
 PN EP832974-A2.
 PD 01-APR-1998.
 PF 11-SEP-1997; 97EP-0115801.
 PR 19-SEP-1996; 96EP-0115001.
 PA (HOFF) HOFFMANN LA ROCHE & CO AG F.
 PI Asakura A, Hoshino T, Ojima S, Shinjoh M, Tomiyama N;
 DR WPI: 1998-195228/18.
 DR N-PSDB: AAV29052.
 XX
 PT Recombinant Gluconobacter oxydans alcohol and/or aldehyde
 PT dehydrogenase enzyme(s) - useful for converting substrate(s), e.g.
 XX L-sorbose or D-sorbitol to 2-keto-L-gulononic acid
 PS Claim 1; Pages 38-40; 59pp; English.
 XX
 CC This is the amino acid sequence for the Gluconobacter oxydans
 CC alcohol and/or aldehyde dehydrogenase A' enzyme. The enzymes
 CC or recombinant organisms can be used to convert suitable substrates
 CC to aldehydes, ketones or carboxylic acids, especially to convert
 CC L-sorbose or D-sorbitol to 2-keto-L-gulononic acid, which can be
 CC converted to L-ascorbic acid by standard procedures. The derivatives
 CC of ADH enzymes have desired substrate specificity, higher affinity
 CC to a substrate, lower affinity to an inhibitory compound, higher
 CC stability against temperature and/or pH and higher catalytic speed.
 XX
 SQ Sequence 579 AA;

Query Match 86.7%; Score 2344.5; DB 19; Length 579;
 Best Local Similarity 78.8%; Pred. No. 1.3e-158;
 Matches 435; Conservative 32; Mismatches 34; Indels 51; Gaps 1;

OY 1 MKPTSLIMASAGALALAAPAQAQVPTDDELANPAGEMISYQONQENRHSPLTQIT 60
 DB 1 mkptslimasagalalAAPAQAQVPTDDELANPAGEMISYQONQENRHSPLTQIT 60
 OY 61 TENVGQLOLVMAKMGKQPKVQVTPPLIHGQVWYLANPGVQIADAKTGLIWEHRQLPN 120
 DB 61 tenvgqlqlvwarqmgpkvqvtpplihdgvywylanpgvdiqaidaktgdlwehrrqlpn 120
 OY 121 IATLNSFG-----I 129
 DB 121 iatlnsfgeptrmalylgtvnyfsvwdhnlvaldtatgvtfdvdrqggedmvsnsqpi 180
 OY 130 VANGYVAGSTCOYSPFCEVSGHDSATGEBELRNRFIPRAGEGDEMGNYEARMTG 189
 DB 130 vangyivagstcoyspfcevsghdsatgelnrnfipragegdegmngnyearmtg 240
 OY 181 vanyivagstcoyspfcevsghdsatgelnrnfipragegdegmngnyearmtg 240
 OY 190 AMGOITPDTNLVHGSTAVGASSETQGTGEGTLYGNTREAVRPDTGETVWRHQTLP 249
 DB 241 amgiltvdpdtnlvhgstavgsasetqgtggtgelygntrfavrpdtgetvwrhqlp 300

QY 250 RDNMDQCTFEEMAVTNVDOPSTMEGLQSTINPNAATGERRVLTGVCKTGIMQFDEAT 309
 XX ||||||| : : : : : ||||||| : : : : : ||||||| : : : : : ||||||| : : : : :
 DB 301 rdnwdeqctfemavtnvdopstmeqlqstlnpnaatgerrvltgvpcktgimqfdaet 360
 QY 310 GEFLARNTNOMMESIDENGIIVTNEDAILKELDYEDVCPFTFLGGRNPSAALNP 369
 ||||||| : : : : : ||||||| : : : : : ||||||| : : : : : ||||||| : : : : :
 DB 361 geftwardtsyenllesidengivtnedailkeldydvcpftflggrnpsaalnp 420
 QY 370 GIYFIPLNNVCYDMAVDOFTSMQVYNTSWTKLPCKDMIGRIDADISTGRTIMSV 429
 ||||||| : : : : : ||||||| : : : : : ||||||| : : : : : ||||||| : : : : :
 DB 421 giyfiplnnvcymavdoftsmqvynstkwtklpckdmigrdadistgtrtlmsv 480
 QY 430 RAAANYSPIVSTGSGVLENGTDRYFRALSOETGETIMOTRLATVYASQAISYEDCMOY 489
 ||||||| : : : : : ||||||| : : : : : ||||||| : : : : : ||||||| : : : : :
 DB 481 reasnypvstlsgavvlfngtdryfralsgetglwqtrlatvaysgavyseldgvy 540
 QY 490 VAIAGGVSYGS 501
 DB 541 vaiaggvstgys 552

RESULT 3

AAW37876 AAW37876 standard; Protein: 579 AA.

AAW37876; (first entry)

10-AUG-1998 (first entry)

Alcohol and/or aldehyde dehydrogenase B amino acid sequence,

Alcohol/aldehyde dehydrogenase B enzyme; recombinant organism;

aldehyde; ketone; carboxylic acid; L-sorbose; D-sorbitol;

2-keto-L-gulonic acid; L-ascorbic; inhibition.

Glucobacter oxydans.

Location/Qualifiers

1..23 /note= "signal peptide"

24..579 /note= "mature protein"

EP832974-A2.

01-APR-1998.

11-SEP-1997; 97EP-0115801.

19-SEP-1996; 96EP-0115001.

(HOPE) HOFFMANN LA ROCHE & CO AG F.

Asakura A, Hoshino T, Ojima S, Shinjoh M, Tomiyama N;

WPI; 1998-195228/18.

N-PSDB: AAV29034.

Recombinant glucobacter oxydans alcohol and/or aldehyde

dehydrogenase enzyme(s) - useful for converting substrate(s), e.g.

L-sorbose or D-sorbitol to 2-keto-L-gulonic acid

Claim 1; Pages 44-46; 59pp; English.

This is the amino acid sequence for the Glucobacter oxydans

alcohol and/or aldehyde dehydrogenase B enzyme. The enzymes

or recombinant organisms can be used to convert suitable substrates

to aldehydes, ketones or carboxylic acids, especially to convert

L-sorbose or D-sorbitol to 2-keto-L-gulonic acid, which can be

converted to L-ascorbic acid by standard procedures. The derivatives

of AABH enzymes have desired substrate specificity, higher affinity

to a substrate, lower affinity to an inhibitory compound, higher

CC stability against temperature and/or pH and higher catalytic speed.

XX Sequence 579 AA;

SO

Query Match 85.4%; Score 2310; DB 19; Length 579;

Best Local Similarity 76.5%; Pred. No. 1,5e-195;

Matches 426; Conservative 38; Mismatches 41; Indels 52; Gaps 1;

1 MKPTSLMASAGLALLAPAFQVPTVDELAPNPAGGWSYQONQENTRHSPLQIT 60

1 mnpflllrtsaavllltpaataqvpicdelanppagewinygrnqenyrhsplqit 60

61 TENYGOLOLWAKMOPKQVQVPLIHGVMTYANQDVYIOADKTDGLTWEHRQLPN 120

61 adnvgvqlvvaigmeagavgvpmldhgvmylanpgdvlaqdaqltlvehrrqlpa 120

121 IATLNSFG----- 128

121 valtnaagdkrkryvalygtstlyfsswdnhlialdmetgvtldvergsgedglsnttgp 180

129 IVANGVIVAGSTCOYSPGCFVSGHDSATGHELMRNTFRPAGBEGDETMGNDYFARMNT 188

181 ivangvivaagstcgyfsgcfvsgahdsatgheilmrntfrpaggdegtmgnfyarmnt 240

189 GAWGQITTDPTNLVHGSFANGPASEFQGTGTYGNTREAVRPTGEIWRKQTL 248

241 gvwgqittdpvtnlvhygsfngpasefqtgttygntreavrpdtgeiwrkqtl 300

249 PRDNMDQCTFEEMAVTNVDOPSTMEGLQSTINPNAATGERRVLTGVCKTGIMQFDEAT 308

301 prdnwdeqctfemavtnvdopstmeqlqstlnpnaatgerrvltgvpcktgimqfdaet 360

309 TGEFLARNTNOMMESIDENGIIVTNEDAILKELDYEDVCPFTFLGGRNPSAALNP 368

361 tgeflarntnomyesidengivtnedailkeldydvcpftflggrnpsaalnp 420

369 SGIFYFIPLNNVCYDMAVDOFTSMQVYNTSWTKLPCKDMIGRIDADISTGRTIMSV 428

421 sgifyfiplnnvcymavdoftsmqvynstkwtklpckdmigrdadistgtrtlmsv 480

429 ERAANYSPIVSTGSGVLENGTDRYFRALSOETGETIMOTRLATVYASQAISYEDCMOY 488

481 erpaanyspivstlsgavvlfngtdryfralsgetglwqtrlatvaysgavyseldgvy 540

489 YVAIAGGVSYGSGLNS 505

541 yvaiaggvstgsglins 557

EP832974-A2.

01-APR-1998.

11-SEP-1997; 97EP-0115801.

19-SEP-1996; 96EP-0115001.

(HOPE) HOFFMANN LA ROCHE & CO AG F.

Asakura A, Hoshino T, Ojima S, Shinjoh M, Tomiyama N;

WPI; 1998-195228/18.

N-PSDB: AAV29034.

Recombinant glucobacter oxydans alcohol and/or aldehyde

dehydrogenase A' enzyme; recombinant organism;

aldehyde; ketone; carboxylic acid; L-sorbose; D-sorbitol;

2-keto-L-gulonic acid; L-ascorbic; inhibition.

Glucobacter oxydans.

Location/Qualifiers

1..23 /note= "signal peptide"

24..578 /note= "mature protein"

EP832974-A2.

XX 01-APR-1998.
 PD 11-SEP-1997; 97EP-0115801.
 PF 19-SEP-1996; 96EP-0115001.
 PR (HOFF) HOFFMANN LA ROCHE & CO AG F.
 PA Asakura A, Hoshino T, Ojima S, Shinjoh M, Tomiyama N.
 PI WPI; 1998-195228/18.
 DR N-PSDB; AAV29053.
 XX
 XX Recombinant Gluconobacter oxydans alcohol and/or aldehyde
 PT dehydrogenase enzyme(s) - useful for converting substrate(s), e.g.
 CC L-sorbose or D-sorbitol to 2-keto-L-gulononic acid
 CC Claim 1; Pages 41-43; 59pp; English.
 PS
 XX This is the amino acid sequence for the Gluconobacter oxydans
 CC alcohol and/or aldehyde dehydrogenase A' enzyme. The enzymes
 CC to recombinant organisms can be used to convert suitable substrates
 CC to aldehydes, ketones or carboxylic acids, especially to convert
 CC L-sorbose or D-sorbitol to 2-keto-L-gulononic acid, which can be
 CC converted to L-ascorbic acid by standard procedures. The derivatives
 CC of AADH enzymes have desired substrate specificity, higher affinity
 CC to a substrate, lower affinity to an inhibitory compound, higher
 CC stability against temperature and/or pH and higher catalytic speed.
 CC
 CC Sequence 578 AA;

Query Match 83.2%; Score 2251; DB 19; Length 578;
 Best Local Similarity 75.7%; Pred. No. 2,5e-190;
 Matches 420; Conservative 37; Mismatches 46; Indels 52; Gaps 2;
 QY 1 MKPSILMASAGALALIAAPAFQVPTVDELAMPAGEWISYQONENRHSPLQIT 60
 DB 1 mktllqssaallvlgilpalag-taitdemlampagewinygqenyrhplcgit 59
 QY 61 TENVGOLQIWMARGOMPKQVYTPHIDGVMTLANPGVIALDAKTGDLWEHRROLPN 120
 DB 60 adhvqglvltwargmeagkivtrpvhqgmymylanpgvdlqalaaacgdlwehrqlpn 119
 QY 121 IATINSFG-----I 129
 DB 120 iatlnsfgeptlgmalylgtlnvfyvswdnhlvaldstsgqvfdvdrgqgtdmvsnsaprl 179
 QY 130 VANGIVINAGSTCOYSPFCFVSGHDSATNGEELMNYRIPRAGEBDEFTMGNDYEARMWTG 189
 DB 180 vanyivagstcgyqpficfsgndsatgeeltrnmflpragegdetvngndyearmwtg 239
 QY 190 AAGGTTTPYNNLVHKGSTAVGRASEMORQKTPGGTLYGTTRAVRPPDGEIWMRQITP 249
 DB 240 wvgqtltpbpyvgvlvhygtsavgrapeaeiqrtvgssmygnittrfavrpelgclwvhtqclp 299
 QY 250 RDNDQECTEBMAYTNVDOPTSTEMEGLSINRNATGRRVLTGVPCKTGTMQFDEAET 309
 DB 300 rdndqectebmnyvndvpsaemegllhainpdaatgerivrtvgpcknglmwgfdaet 359
 QY 310 GEFLMADITVYOMMIESIDENGIVYNEDAILKELDVEVDVCPPTLGRDMPALNDS 369
 DB 360 geflwaditcynlliesvdpdglvhnedlvtelevayeloptligrfdwpsaallpndt 419
 QY 370 GIYFIPLNVCYDMAVDOFTSMOYNTSNTKLPKPKMDIGRIDALIDISTGRTLSVE 429
 DB 420 giylfiplnacsgmtavdefgsldvnyvsldkispsemgrtldidistgltlswae 479
 QY 430 RAALANSPVLTSGGVLFWGSDTRFALSOETGETLMOQLATLAVASGOALSTEVGMQY 489
 DB 480 ryasnfpavltsggvlfwngsdtrfalsgetgetlmoqlatlasvasgaisylsdvqy 539

QY 490 VAIAGGVSVXSGSLN 504
 DB 540 valigrvtsgslshn 554

RESULT 5
 ID AAB35987 standard; Protein; 754 AA.
 AC AAB35987;
 DF 01-MAR-2001 (first entry)
 DE Sorbitol dehydrogenase subunit 1 amino acid sequence.
 XX Sorbitol dehydrogenase; SDH; Gluconobacter oxydans; D-sorbitol;
 KW L-sorbose production; 2-keto-L-gulononic acid.
 XX
 XX Gluconobacter oxydans.
 OS
 XX WO2000065066-A1.
 PN
 XX 02-NOV-2000.
 PD
 XX 23-APR-1999; 99WO-1B00736.
 PF
 XX 23-APR-1999; 99WO-1B00736.
 PR
 XX 23-APR-1999; 99WO-1B00736.
 PA (CHOI/) CHOI E.
 PA (RHEE/) RHEE S.
 PA (LEE/) LEE E.
 XX
 PI Choi E, Rhee S, Lee E;
 XX
 DR WPI; 2000-687351/67.
 DR N-PSDB; AAC83153.

Novel membrane-bound sorbitol dehydrogenase nucleic acid molecule isolated from Gluconobacter suboxydans useful for the fermentative production of 2-keto-L-gulononic acid and L-sorbose from D-sorbitol - Claim 1; Fig 8; 96pp; English.
 CC This invention relates to an isolated membrane-bound sorbitol
 CC dehydrogenase (SDH) from Gluconobacter oxydans. The invention includes
 CC polynucleotides (AAC83153 - AAC83154) encoding the 3 subunits of SDH
 CC (AAB35987 - AAB35989). Also included in the invention are two
 CC polynucleotide sequences AAC83156 and AAC83157 which contain the subunit
 CC coding regions. The sorbitol dehydrogenase polynucleotide sequences are
 CC useful for producing L-sorbose from D-sorbitol and for increasing the
 CC production of 2-keto-L-gulononic acid by transforming a host cell.
 CC especially Gluconobacter with the DNA and selecting the transformed host
 CC genes) encoding fragments of SDH sequences (fragments of the SDH subunit
 CC sequences are not given separately in the specification but are included
 CC in sequences AAC83156 and AAC83157.
 CC
 CC Sequence 754 AA;

Query Match 17.8%; Score 480.5; DB 21; Length 754;
 Best Local Similarity 26.2%; Pred. No. 1.8e-33;
 Matches 157; Conservative 80; Mismatches 234; Indels 129; Gaps 17;
 QY 14 LALLAAPAFQVPTV-----TDELAMPAGEWISYQONENRHSPLQITTENVGQ 66
 DB 17 lgcataalafcatpsvalaedtgatlnedqhpqdwmsygrtysqrlpqlgikdhasn 76
 QY 67 LQIWMARGMPKAVQV-TPLIHDGVMTLANPGDVIOALDAKTGDLWEHRROLPN 122
 DB 77 lklawhydlatnrggeqtlpivdyvmyatlnwskmkaldaatgllvsvdpkypgnladr 136
 QY 123 -----TINSFGIVANGIVAGS-----TCQYS 144

10 SAGLALLAMPARROTPTTDE-----LLANPAGEISVYSGOENOEHRHSPLTQIT 61
17 tagliccaallisyatamasadggatgaallihaddpgrmwtgrtytdsqtrpsldqlnr 76
62 ENNGCOLVWARKOMPGRKQV-TPILINDGMYWLANGDVIOALIDKTCGDLIWEHRRLP- 119
77 snygnlkawylldltnrgsgtrplvldgymwattmnmkavdaatgkllwsydpvrg 136
120 NIA-----TLNS-----FGI----- 129
137 niadgscctvrrgaayngkvfyfgrlialadlaktgklwagsvntllpceaalqkrsy 196
130 -----VANGVIYASTCOYSPFGC--TVSGHDSATGEELMRNYFTPRAGEGD----- 175
197 tvdgprlakgrvllgn--gsefsggrtvasafdeitgkvawrtflvprnpkhepdaas 254
176 -----ETMGNDYEARNMGA-----WGOITDPVTNLVHGSLTAGAPASEQR 218
255 vlmnkaytws-----ptgawttrggggrtwwdsiypdvadlvlylgvngsppnryyr 307
219 GRPGTILXGINTRANRANPDGETELWRHDTLPDMMDOCEFFEMMYNVNVDPSTMBESLO 278
308 seegkydnlfigslvaalkpetgryvwhfgetpmgdwtfsdqmlclpI----- 357
279 SINPNAATGERRVLTVCPRKCTGMQOPAEFGEELMARDTNNOMMIESID-ENGIVRYNE 337
358 -----ngcthrivvthangfyfllaangetiesgkyvyyvaasglpdkigrpIynp 410
338 DALIKELDVYDPCPTFLGGRDWPASALINPDGIFYETPIPNNVCYMMMAVDEEFT-SNDVY 396
411 dalytltgkewygiipgdldgnhfaamafspkigtlyiypqgvpilytqvggftipbds 470
397 NTS---NVKRLPPG-----KDMIRIDALIDISTGRTLMSVERAANYSPVLTSGGCV 445
471 nlglmknkvipdspeakqatvklklywvawpdpqgaearvrdhkpwnvgjllaaggl 530
446 LFNNGDTRFRALSOEGETIMQTRLATYVAAAGSAISEYEDMOVATVIAAG 495
531 lftglangefnaydalngsdlnfhaadsllipptylangkqyvvaaveng 580

[illegible]

[illegible]

RESULT	10
ABG24430	
ID	ABG24430 standard; Protein; 443 AA.
XX	
AC	ABG24430;
XX	
DT	18-FEB-2002 (first entry)
XX	
DE	Novel human diagnostic protein #24421.
XX	
KW	Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX	food supplement; medical imaging; diagnostic; genetic disorder.
KW	
XX	
OS	Homo sapiens.
XX	
PN	W0200175067-A2.
XX	
PD	11-OCT-2001.
XX	

PF 30-MAR-2001; 2001WO-US08631.
 XX
 PR 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 XX
 PA (HYSE-) HYSEQ INC.

PI Dmanac RT, Liu C, Tang YF;
 XX WPI: 2001-639362/73.
 DR N-PSDB; AAS88617.
 XX

PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 XX biodiversity

PS Claim 20; SEQ ID No 54789; 103bp; English.

XX The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations in
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. AAG00010-ABG30377 represent novel human
 CC diagnostic amino acid sequences. AAG00010-ABG30377 represent novel human
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pat_sequences.
 XX
 SO Sequence 443 AA;

Query Match 5.7%; Score 155.5; DB 22; Length 443;
 Best Local Similarity 31.5%; Pred. No. 4.7e-05;
 Matches 46; Conservative 21; Mismatches 50; Indels 29; Gaps 7;

OY 6 LTMAS-----AGALALAAPFAOVTPVTDELLANPPAGEWISYGNONENYRHSPLTQ 58
 DB 168 ltwagfndpgeingltsadatpaea-ispvadt-----dwpaygrngqgqitfslkq 218
 OY 59 ITTENVGOLQVMA-----RCMGPGKV--GVTPLIHDGVWYLANEGDVIOAIDAKTGD 109
 DB 219 inadhvnhlkeawftrgdvkgpndpbeltnetvtpikvgtclctahqylfaidaasgk 278
 OY 110 LIMEHRQLNINATLNSF-GIVANGV 134
 DB 279 ekwhyd---pelktnesighvtrcgv 301

RESULT 11
 ID AAM09614 standard; Protein: 824 AA.
 AC AAM09614;
 XX
 DT 20-MAY-1997 (first entry)
 XX
 DE Pathogenic Staphylococcus aureus ORF1 gene product.
 XX
 KW Pathogenicity; vaccine; immunoassay; antibody.
 XX

OS Staphylococcus aureus strain NCTC 8325-4 (ISP8).
 XX
 PN W09641878-A1.
 XX
 PD 27-DEC-1996.
 XX

PF 13-JUN-1996; 96WO-AU00353.
 XX
 PR 13-JUN-1995; 95AU-0003507.
 XX

PA (AUSU) UNIV AUSTRALIAN NAT.
 XX
 PI Adham El-Adhami W, Matthei KI, Stewart PR.
 XX
 DR WPI: 1997-065465/06.
 DR N-PSDB; AAT47757.
 XX

PT Nucleic acids specific to pathogenic Staphylococcus aureus strains
 PT for diagnosis and immunogenic protein prodn.
 XX
 PS Claim 30; Page 62-65; 88bp; English.

XX An 824-amino acid protein (AAM09614) and 192-amino acid protein
 CC (AAM09615) were identified as products of 2 open reading frames in
 CC a nucleic acid (AAT47757) isolated from pathogenic Staphylococcus
 CC aureus strain ISP8. The 2 gene products are highly hydrophilic
 CC and have high surface probability and antigenic index measurements,
 CC suggesting that they are surface expressed. The polypeptides,
 CC vaccines and as immunoassay reagents for detecting pathogenic S.
 CC aureus strains.

SO Sequence 824 AA;

Query Match 5.0%; Score 134.5; DB 18; Length 824;
 Best Local Similarity 19.3%; Pred. No. 0.0087;
 Matches 88; Conservative 58; Mismatches 131; Indels 179; Gaps 20;

OY 19 APAFAOVTPVTDELLANPPAGEWISYGNONENYRHSPLTQITTEVNGOLQV 70
 DB 427 apglekviregqkgektlptlknpllgeliskgeskeeltkdpinel----- 476
 OY 71 WARGMOPGVVPLIHD-----GVWYLANPGDVIOAIDAKTGDPLIWEHR 115
 DB 477 -----eygpetlcpghndefdklptgekeevpkygkkn-----etgdv----- 518
 OY 116 ROLPNITLNSFETIVANGVIVAGSTCOYSPFCFVSGHDSATGELMRNFIFRAGEED 175
 DB 519 -----fppvdsyqkygypvkgdsivekeelpfkkektf--npdlapgtex-----viregkge 569
 OY 176 ETWGNDEYEARWMTG-----AMGQITDPTNIVHVGSTAVGPA-----SE 215
 DB 570 ktltptlknpllgeliskgeskeeltkdpinelteygpelcpghndefdklptgeke 629
 OY 216 TORPGGTGLXTNTRFAVRPT-----GEIYWRHQTLPDRNMOEOTFEAMATN 265
 DB 630 evpkykpy--lknpekgdvrvpvdsvckypvkgdsivekeelp-----fekerktf-----n 679
 OY 266 VDVQPSST-----MEGLQSI----- 280
 DB 680 pdlapgtelkviregqkgektlptlknpllgeliskgeskeeltkdpinelteygpel 739
 OY 281 -----NPNMATGERRVLTGVPCRTGTGMQFDEATGEGFLIARNTNNONNIESIDNGI 332
 DB 740 tprhndefdklptgeke-----evpkygkik--npektgdv-----trpvdsvckyg- 785
 OY 333 VVNEDEILKELDEYDVCTPTFLGGRDMSALNDP 368
 DB 786 -pvkgdsivekeelpf-----kerktfnd 809

RESULT 12

AA681738
ID AAG81738 standard; Protein: 715 AA.
XX
XX AAG81738:
XX 03-SEP-2001 (first entry)
XX
XX S. epidermidis open reading frame protein sequence SEQ ID NO:570.
XX
XX Staphylococcus epidermidis SRI strain; infection; diagnosis;
XX
XX Staphylococcus epidermidis.
XX
XX Staphylococcus epidermidis.
XX
XX W0200134809-A2.
XX
XX 17-MAY-2001.
XX
XX 09-NOV-2000; 2000MO-US30782.
XX
XX 09-NOV-1999; 99US-0164258.
XX
XX (GLAX) GLAXO GROUP LTD.
XX
XX Kimmerly WJ;
XX
XX WPI: 2001-316495/33.
XX
XX N-PSDB: AAH52388.
XX
XX Nucleic acids encoding polypeptides from Staphylococcus epidermidis,
XX
XX useful for vaccinating against infections, e.g. endocarditis -
XX
XX Claim 18; Page 189; 2188pp: English.

XX
XX AAH53970 represent nucleic acids (I) encoding polypeptides
XX
XX (II), given in AAG81454 to AAG83120, from Staphylococcus epidermidis.
XX
XX (I) and (II) can have antibacterial activity and therefore can be used
XX
XX in vaccination. The nucleic acids (I) may be used to produce the
XX
XX S. epidermidis polypeptides (II) via the production of vectors
XX
XX containing them which are used to produce hosts cells which express the
XX
XX polypeptides. The polypeptides (II) (and/or nucleic acids) may then be
XX
XX used to vaccinate subjects and to raise antibodies against the bacteria.
XX
XX The polypeptides may also be used to assay for other inhibitors of their
XX
XX activity and therefore identify compounds that may be used for the
XX
XX treatment of S. epidermidis infections, e.g. endocarditis. AAH53971 to
XX
XX AAH55090 represent specifically claimed S. epidermidis genomic DNA
XX
XX polynucleotide sequences from the present invention. AAH55091 to
XX
XX AAH55098 represent oligonucleotide sequences and primers which are used
XX
XX in the exemplification of the present invention.
XX
XX N.B. The present invention specifically claims all the polynucleotide
XX
XX sequences given in the sequence listing of the present specification,
XX
XX however the sequence listing only goes up to SEQ ID NO:4454 so even
XX
XX though sequences are given in the disclosure for SEQ ID NO:4465 to 4472,
XX
XX no sequences are present for SEQ ID NO:4455 to 4464.

XX
XX Sequence 715 AA;

Query Match 5.0%: Score 134; DB 22; Length 715;
Best Local Similarity 23.5%; Pred. No. 0.0076;
Matches 104; Conservative 52; Mismatches 160; Indels 126; Gaps 26;

QY 26 TPTVDELLANPAGEMISYQONENYRHSPLTQIT-----TENVGQQLVYMAAGMO- 76
DB 309 tptt-----knptlgekgvgeptekvtkpydeivhyggeekpghdeldfnapksqg 364
QY 77 --DGR-----VQYPLIHDGVMYLANPGDVIOAIDAKTGDLIWEHRRQL-PNITATL 124
DB 365 dvgkpgvknptdgvvrrpvdvktkypvdgdp!-----tsteipdkkrefpnlap- 419
QY 125 NSGIVANGVYAGSTCGYSPGCEVSGHDSATGSEELMRNTFITRAGEEGDEGTMGNDYEA 184
DB 420 gtekvvgkg--epgtktlttpt-----ttknptlgekv-----gegept----- 455

QY 185 RMTGANGQITTPYVNLVHYGSTAYGP-----ASETQRTPGCTLLGTNTR 231
DB 456 -----ekvltkpydeivhyggeekpghdeldfnapksqgttkpkyg----- 499
QY 232 FAVRPDGEIYWRHQTLPDRNWDCECTFEMWTVN---VVOVSTMEGL-----OSINPN 283
DB 500 -vknpdigev-----tpvdd-----vlykypvdgdptlsteelpfkkrfnpd 544
QY 284 AATGERRV-LTGVP-CKT-GTMMQFPAETGEFTWARDTNYQNMIESIDENGIYVNDAT 340
DB 545 lkpgeerwkqkgepytklttptknlptkgekgvgegepektkqpvde--lleyggee 602
QY 341 LKELDEVDCPTFLGGRWPS--AALNPDG--YFIPLNNVCY-----DMMAYDOR-- 389
DB 603 kpgkhdefdnapksqgdvpgkpvknptdgvvrrpvdvktkypvdgdptlsteelp 662
QY 390 FTSMDVYNTSNVTKLPKPKDMI 411
DB 663 fdkkrefnpd---lkpkerv 680

RESULT 13

AA683007
ID AAG83007 standard; Protein: 910 AA.
XX
XX AAG83007:
XX
XX 03-SEP-2001 (first entry)
XX
XX S. epidermidis open reading frame protein sequence SEQ ID NO:3108.
XX
XX Staphylococcus epidermidis SRI strain; infection; diagnosis;
XX
XX Staphylococcus epidermidis.
XX
XX Staphylococcus epidermidis.
XX
XX Staphylococcus epidermidis.
XX
XX W0200134809-A2.
XX
XX 17-MAY-2001.
XX
XX 09-NOV-2000; 2000MO-US30782.
XX
XX 09-NOV-1999; 99US-0164258.
XX
XX (GLAX) GLAXO GROUP LTD.
XX
XX Kimmerly WJ;
XX
XX WPI: 2001-316495/33.
XX
XX N-PSDB: AAH53857.
XX
XX Nucleic acids encoding polypeptides from Staphylococcus epidermidis,
XX
XX useful for vaccinating against infections, e.g. endocarditis -
XX
XX Claim 18; Page 819; 2188pp: English.

XX
XX AAH53970 represent nucleic acids (I) encoding polypeptides
XX
XX (II), given in AAG81454 to AAG83120, from Staphylococcus epidermidis.
XX
XX (I) and (II) can have antibacterial activity and therefore can be used
XX
XX in vaccination. The nucleic acids (I) may be used to produce the
XX
XX S. epidermidis polypeptides (II) via the production of vectors
XX
XX containing them which are used to produce hosts cells which express the
XX
XX polypeptides. The polypeptides (II) (and/or nucleic acids) may then be
XX
XX used to vaccinate subjects and to raise antibodies against the bacteria.
XX
XX The polypeptides may also be used to assay for other inhibitors of their
XX
XX activity and therefore identify compounds that may be used for the
XX
XX treatment of S. epidermidis infections, e.g. endocarditis. AAH53971 to
XX
XX AAH55090 represent specifically claimed S. epidermidis genomic DNA
XX
XX polynucleotide sequences from the present invention. AAH55091 to
XX
XX AAH55098 represent oligonucleotide sequences and primers which are used
XX
XX in the exemplification of the present invention.
XX
XX N.B. The present invention specifically claims all the polynucleotide

CC sequences given in the sequence listing of the present specification,
 CC however the sequence listing only goes up to SEQ ID NO:4454 so even
 CC though sequences are given in the disclosure for SEQ ID NO:4465 to 4472,
 CC no sequences are present for SEQ ID NO:4455 to 4464.
 XX
 XX Sequence 910 AA;

Query Match 4.9%; Score 133.5; DB 22; Length 910;
 Best Local Similarity 25.1%; Pred. No. 0.012;
 Matches 97; Conservative 46; Mismatches 151; Indels 93; Gaps 23;

QY 26 TPVTDELLANPAGEMWISYQONENYRHSPLQITTT---ENV-----GQLQVWARGMQ- 76
 DB 231 tptt-----knplgtkvgegeptektkpvdeltgysgeelkpgkhdefpnabkysge 286
 QY 77 --GK-----VOVPLIHGVMYLANPQVYIAIDAKTGDLIWEHROL-PNIA TL 124
 DB 287 dvpgkpyvknptdgvvtpvddvdkypvddp-----tsteelpfdkkrtdfnlap- 341
 QY 125 NSFQIVANGVIYVAGSCQSPFCVSGHDSATGEELMRNYFIPRAGEDEDTWGNDEYA 184
 DB 342 gtekvvgkg--epgkltltp-----tknplgtkv-----gegept----- 377
 QY 185 RMTGAMGOIYDPTNIVHGSTAVGPA--SETQRGTPGTLGCTNTRAVR-PDTGEI 241
 DB 378 -----ekvtkpyvdelvhygeelkpgkhdefdnabkysgedvpgkpyvknptdgv 430
 QY 242 VMRHQTLPDNDMDCECTEFEMAVTN---VDVOPSTEMEGL-----QSTPNAATG-ERRYL 292
 DB 431 v-----tpvdd-----vltkypvddgdplsteeelpfdkkrtdfnlapgtekvvq 476
 QY 293 TGVF-CGT-GTMMQFDETFEFLMARDTNYQNMIESIDENGITVVEDAILKELDERYDV 350
 DB 477 kgpgtkltltpktnplgtkvgegeptektkpvdv-----lvhyggeelkpgkhdefdp 534
 QY 351 CPTFLGGRWMS--AALNPSGTYFIP 375
 DB 535 napkysgedvpgkpyvknptdgvvtp 561

RESULT 14

AAG82914
 ID AAG82914 standard; Protein: 696 AA.

XX AC AAG82914;
 XX DT 03-SEP-2001 (first entry)
 XX DE S. epidermidis open reading frame protein sequence SEQ ID NO:2922.
 XX KW Staphylococcus epidermidis SRI strain; infection; diagnosis;
 XX OS Staphylococcus epidermidis.
 XX PN WO200134809-A2.
 XX PD 17-MAY-2001.
 XX PF 09-NOV-2000; 2000WO-US30782.
 XX PR 09-NOV-1999; 9905-0164258.
 XX PA (GLAX) GLAXO GROUP LTD.
 XX PI Kimmery WJ;
 XX DR WPI: 2001-316495/33.
 XX N-PSDB: AAH53764.
 PT Nucleic acids encoding polypeptides from Staphylococcus epidermidis,
 PT useful for vaccinating against infections, e.g. endocarditis -

XX
 PS Claim 18; Page 763-764; 2188bp; English.

XX AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides
 CC (II), given in AAG81454 to AAG83120, from Staphylococcus epidermidis.
 CC (I) and (II) can have antibacterial activity and therefore can be used
 CC in vaccination. The nucleic acids (I) may be used to produce the
 CC S. epidermidis polypeptides (II) via the production of vectors
 CC containing them which are used to produce hosts cells which express the
 CC polypeptides. The polypeptides (II) (and/or nucleic acids) may then be
 CC used to vaccinate subjects and to raise antibodies against the bacteria.
 CC activity and therefore identify compounds that may be used for the
 CC treatment of S. epidermidis infections, e.g. endocarditis. AAH53971 to
 CC AAH55090 represent specifically claimed S. epidermidis genomic DNA
 CC AAH55098 represent oligonucleotide sequences from the present invention.
 CC N. B. The present invention specifically claims all the polynucleotide
 CC sequences given in the sequence listing of the present specification,
 CC however the sequence listing only goes up to SEQ ID NO:4454 so even
 CC though sequences are given in the disclosure for SEQ ID NO:4465 to 4472,
 CC no sequences are present for SEQ ID NO:4455 to 4464.
 XX
 XX Sequence 696 AA;

Query Match 4.9%; Score 133; DB 22; Length 696;
 Best Local Similarity 23.9%; Pred. No. 0.0092;
 Matches 106; Conservative 57; Mismatches 152; Indels 128; Gaps 27;

QY 26 TPVTDELLANPAGEMWISYQONENYRHSPLQITTT---ENV-----GQLQVWARGMQ- 76
 DB 103 tptt-----knplgtkvgegeptektkpvdeltgysgeelkpgkhdefdnabkysge 158
 QY 77 --GK-----VOVPLIHGVMYLANPQVYIAIDAKTGDLIWEHROL-PNIA TL 124
 DB 159 dvpgkpyvknptdgvvtpvddvdkypvddp-----tsteelpfdkkrtdfnlap- 213
 QY 125 NSFQIVANGVIYVAGSCQSPFCVSGHDSATGEELMRNYFIPRAGEDEDTWGNDEYA 184
 DB 214 gtekvvgkg--epgkltltp-----tknplgtkv-----gegept----- 249
 QY 185 RMTGAMGOIYDPTNIVHGSTAVGPA--SETQRGTPGTLGCTNTRAVR-PDTGEI 241
 DB 250 -----ekltkpyvdelvhygeelkpgkhdefdnabkysgedvpgkpyvknptdgv 302
 QY 242 VMRHQTLPDNDMDCECTEFEMAVTN---VDVOPSTEMEGL-----QSTPNAATG-ERRYL 292
 DB 303 v-----tpvdd-----vltkypvddgdplsteeelpfdkkrtdfnlapgtekvvq 348
 QY 293 TGVF-CGT-GTMMQFDETFEFLMARDTNYQNMIESIDENGITVVEDAILKELDERYDV 350
 DB 349 kgpgtkltltpktnplgtkvgegeptektkpvdv-----lvhyggeelkpgkhdefdp 405
 QY 351 CPTFLGGRWMSAAL-----NPSG--IYFIPLNWCYDMMAVDOEFTSMD-- 394
 DB 406 -----pnapsketeypgkpyvknptdgvvtpvddvdkypkvgnplsteeel 455
 QY 395 -----VYNTSNVTKLPQKDMI 411
 DB 456 pfdkkrvfnpd-----lkpggeev 474

RESULT 15

AAU34349
 ID AAU34349 standard; Protein: 1016 AA.

XX AC AAU34349;
 XX DT 14-FEB-2002 (first entry)
 XX DE Staphylococcus aureus cellular proliferation protein #625.

XX Antisense: prokaryotic cellular proliferation protein;
KM antibiotic; antibacterial; drug design.

XX Staphylococcus aureus.

XX WO200170955-A2.

XX 27-SEP-2001.

XX 21-MAR-2001; 2001WO-US09180.

XX 21-MAR-2000; 2000US-191078P.

XX 23-MAY-2000; 2000US-206848P.

XX 26-MAY-2000; 2000US-207727P.

XX 23-OCT-2000; 2000US-242578P.

XX 27-NOV-2000; 2000US-253625P.

XX 22-DEC-2000; 2000US-257931P.

XX 16-FEB-2001; 2001US-269308P.

XX (ELIT-) ELITRA PHARM INC.

XX Haselbeck R, Ohlson KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI Yamamoto RT, Xu HH;
XX WPI: 2001-611495/70.
XX N-PSDB: AAS52208.

XX New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids -
XX

XX Example 3; Seq ID No 5845; 511pp; English.

XX The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the
CC genes, their use in the discovery of novel antibiotics, the essential
CC genes themselves and the encoded proteins. The prokaryotes used are
CC *Escherichia coli*, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella*
CC *pneumoniae*, *Pseudomonas aeruginosa* and *Enterococcus faecalis*. The
CC invention is also useful for the identification of potential new targets
CC for antibiotic development. The antisense nucleic acids can also be used
CC to identify proteins used in proliferation, to express these proteins,
CC and to obtain antibodies capable of binding to the expressed proteins.
CC The proteins can be used to screen compounds in rational drug discovery
CC programmes. The antisense nucleic acid sequence is also useful to screen
CC for homologous nucleic acids which are required for cell proliferation in
CC a wide variety of organisms. The present sequence represents an
CC essential prokaryotic cellular proliferation protein.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 1016 AA;

Query Match 4.9%; Score 133; DB 22; Length 1016;
Best Local Similarity 21.2%; Pred. No. 0.016; Indels 184; Gaps 27;
Matches 113; Conservative 59; Mismatches 176;

OY 19 APAAQV-----PVTDELANPPAGENISYONONENRHSPLTQITTENYGOLQV 70
DB 528 apgtekvtregqkgektltptlknpitgvliskgeskeeltkdpinelt----- 577
OY 71 WARGQPKQVYPLIHD-----GVWYLANPGDVYQAIIDAKTGDLIWEHR 115
DB 578 -----eypetitpghndetfdpkipgekeevpkipnp-----evgdyv----- 619
OY 116 ROLPNIATLNSFGIYAVAGSTCOYSPFGCVSGHSDATGEELWRYFLPRAGCEGD 175
DB 620 --rppdvstkygvpkgsivkeelpfekekt--npdlapptek-----viregkge 670
OY 176 ETWGNDEYKRWMTGA-----WQITTYDPTVTLVHGSTAVGPA-----SE 215

DB 671 kttptlknpitgvliskgepkeeltkdpineltgypetltphndetfdpkipgeke 730
OY 216 TORPAGSTLYGTNTRPAPRPT-----GELVNRHOTLPDNDMDOCSTFEMAVTN 265
DB 731 evpkipg--lknpetgdvtrpypdvskypgvkgsivkeelp--fkertkf-----n 780
OY 266 VVVPSTEMEGLQSLNPAATGERRVLTGVPCRTGTMOPDAETGEFLMARTNYONMIE 325
DB 781 pdlappte-----kviregkgektitc--ptlknp-----tgeit-----skge 819
OY 326 SIDENGIYVNEADAILKELDVEYDVCPFLGGRDWPASAL-----NPSG 370
DB 820 sice-----itkdpinel--teygpeltipghndetfdpkipgekeevpkipnpetcy 872
OY 371 IYIPLANNVCYDMADVDEFTSMDEVNTSNVTKLPK--DMIGRIDALIDSTGRTLSVE 905
DB 873 dvvtrp-----pvd-----svtkypgvkgsivkeelpfeke 905
OY 430 RAAANSPVLSTGGVLFNGGDRFRALSOETGTLWOTRLATVAGSQAIS 481
DB 906 ----kfnplap-----gtekvire-gkgektitptlknpitgvlisk 944

Search completed: May 24, 2002, 10:19:20
Job time: 383 sec

Fri May 24 11:27:28 2002

walick-934-128.pcp.rag

Fri May 24 11:27:29 2002

walick-934-128.pep.ra1

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2002, 10:20:08 ; Search time 39.53 Seconds
(without alignments)
312.040 Million cell updates/sec

Title: WALICK-934-128.PEP

Perfect score: 1 MKPTSLMASAGLALALAP.....GMOYVALAGGVSYSGSLNS 505

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/6C.COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/6D.COMB.pep:*
Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	480.5	17.8	US-09-296-284-4	Sequence 4, Appl
2	479	17.7	US-07-985-458-3	Sequence 3, Appl
3	471.5	17.4	US-09-296-284-25	Sequence 2, Appl
4	319	11.8	US-09-136-251-2	Sequence 2, Appl
5	106	3.9	US-08-868-786-2	Sequence 10, Appl
6	105	3.9	US-08-219-262B-10	Sequence 10, Appl
7	105	3.9	US-09-031-655-10	Sequence 11, Appl
8	103.5	3.8	US-08-868-786-6	Sequence 11, Appl
9	101.5	3.8	US-08-928-361B-11	Sequence 16, Appl
10	101.5	3.8	US-08-928-361B-5	Sequence 16, Appl
11	99.5	3.7	US-09-540-245A-16	Sequence 4, Appl
12	99.5	3.7	US-07-642-734C-4	Sequence 4, Appl
13	99.5	3.7	US-08-439-009A-4	Sequence 4, Appl
14	97	3.6	US-08-856-841-22	Sequence 22, Appl
15	97	3.6	US-08-700-651-5	Sequence 5, Appl
16	96.5	3.6	US-08-883-534-3	Sequence 3, Appl
17	96.5	3.6	US-09-204-764-3	Sequence 3, Appl
18	95.5	3.5	PCR-US93-10541-2	Sequence 2, Appl
19	95.5	3.5	US-08-987-519-2	Sequence 2, Appl
20	95	3.5	US-07-731-157A-4	Sequence 4, Appl
21	95	3.5	US-08-229-444B-2	Sequence 4, Appl
22	95	3.5	US-08-541-780-4	Sequence 19, Appl
23	95	3.5	US-08-216-276A-19	Sequence 2, Appl
24	94.5	3.5	US-08-714-402-2	Sequence 2, Appl
25	94.5	3.5	US-09-356-818A-2	Sequence 10, Appl
26	94.5	3.5	US-07-462-484-10	Sequence 10, Appl
27	94.5	3.5	US-08-441-147-10	Sequence 10, Appl

28	94.5	3.5	527	5	PCR-US95-07536-10	Sequence 10, Appl
29	94	3.5	635	4	US-08-931-608A-5	Sequence 5, Appl
30	94	3.5	824	1	US-09-626-589-3	Sequence 3, Appl
31	93.5	3.5	548	1	US-08-247-902A-2	Sequence 2, Appl
32	93	3.4	380	3	US-08-971-782-4	Sequence 4, Appl
33	93	3.4	380	4	US-09-309-026-4	Sequence 4, Appl
34	93	3.4	459	3	US-08-971-782-2	Sequence 2, Appl
35	92.5	3.4	459	4	US-09-309-026-2	Sequence 2, Appl
36	92.5	3.4	659	4	US-09-626-589-2	Sequence 1, Appl
37	92.5	3.4	717	4	US-09-626-589-1	Sequence 2, Appl
38	92.5	3.4	5215	4	US-09-105-537-2	Sequence 1, Appl
39	91.5	3.4	532	3	US-08-737-336-6	Sequence 2, Appl
40	91.5	3.4	954	4	US-09-251-645-12	Sequence 12, Appl
41	91.5	3.4	1612	1	US-08-169-927-1	Sequence 2, Appl
42	91	3.4	500	2	US-08-987-519-1	Sequence 6, Appl
43	91	3.4	551	2	US-08-793-229-32	Sequence 32, Appl
44	91	3.4	551	3	US-09-285-957-32	Sequence 32, Appl
45	91	3.4	682	3	US-08-481-435-6	Sequence 6, Appl

ALIGNMENTS

RESULT 1
US-09-296-284-4
Sequence 4, Application US/09296284A
Patent No. 6204040
GENERAL INFORMATION:
APPLICANT: Choi, Eun-Sung
APPLICANT: Rhee, Sang-Ki
APPLICANT: Lee, Eun-Hae
TITLE OF INVENTION: Glucosyltransferase Suboxylans Sorbitol Dehydrogenase, Genes
FILE REFERENCE: 1533-0870000
CURRENT APPLICATION NUMBER: US/09/296,284A
CURRENT FILING DATE: 1999-04-22
NUMBER OF SEQ ID NOS: 87
SOFTWARE: Patent Ver. 2.0
SEQ ID NO 4
LENGTH: 754
TYPE: PRT
ORGANISM: Glucosyltransferase suboxylans
US-09-296-284-4

Query Match	17.8%	Score 480.5	DB 4	Length 754
Best Local Similarity	26.2%	Pred. No. 1.5e-37		
Matches 157	Conservative	80	Mismatches 234	Indels 129
Gaps 17				
QY	14	LALLAPPAFAOVTV	-----TDELLAPPAGMEISYQONENRHSPLTQTITENVGQ	66
DB	17	LGCMAALAFCAITSPVALAEDTGTATINADQHGDMMSYGRYSRSPDIOITKDNASN	76	
QY	67	LQVWAROMQGVQV	-----TPIIHGCVWLANPQDVLOAIADAKGDLIMHRRQLP-NIA--	122
DB	77	LKLANHYLDLNRQEGPILVDGVYATTNWSKRALDAATKILMSYDCKVGNIDR	136	
QY	123	-----TINSFGIVANCVIYAGS-----	-----TCOYS	144
DB	137	GCDDVYNRGAAYMNKVFYFTFGDRLTALDAKGLIWSVYTPKRAOLGHRSTYVNGA	196	
QY	145	P-----FGC-----FVSGHDSANGELIMRYEIPRAGEGD-----	-----	175
DB	197	PRIAKKVILIGGAEFGARGVTAIDAETGKMDWRFVTPNDKRPDGAASDQVILMSKA	256	
QY	176	ETWMDYERAMWTG-----AMGQITVPTNLIYHGSTAVPASSETORGTGPGCTLTNTNR	231	
DB	257	YPTWGGAGAKQGGGGGVWDSLITDPTDLVYLGVGNGSPWNYKFRSECKGNLFLGSI	316	
QY	232	FAVAPDTGELIWHHQTLPDNNNDQCTEDMAYTVNDVQSPTEWGLQISINPAATGE-RR	290	
DB	317	VAIMPDTGKTVWHFQFTPMQDMQDYSVOOINALDKPV-----NGEMRH	359	

```

QY 291 VITGVCKTGTMOFDETEGFLMARDTNYONMIESID-NCITVNDALIKELDEYD 349
Db 360 VLVHAP--KNGFYIIDAKTKFTISGRPYTEWMANGLDVGRPNYNDALMTLNKRPWY 418
QY 350 VCFPLGGRMPSALNPDGSIYFIPLNVCY-----DMVAVOQETSNDY 396
Db 419 GIPDGLGHNFAAASPOKTLVYIPAQOVPPVYDPOKGRKHHDSMNLGLDNKKIGL 478
QY 397 NTSN---VTKLPKGRDMIGRIDDAIDISTGRILMSVEEAANYSPLVSTGGVLFNGGTD 452
Db 479 DDNDPOHKADKQOFLDKLGVAMPDPOKQOATFVDBKGFWMGGLATAGVLFQGLAN 538
QY 453 RYFALSQETGETIMOTRLATYASGOAISYEVDGMOYVAI---AG-----GGVSYGSG 502
Db 539 GEHHAADATGKDLFFPPQASAIIPVPTYANGKQYVAEYVGMGIIYFFLGGVARTSG 598

```

RESULT 2

US-07-985-458-3
Sequence 3, Application US/07985458

GENERAL INFORMATION:

APPLICANT: Tamaki, Toshimi;
APPLICANT: Takemura, Hiroshi;
APPLICANT: Takemura, Kenji;
APPLICANT: Fukaya, Masahiro;
APPLICANT: Okumura, Hajime and
APPLICANT: Kawamura, Yoshiya
TITLE OF INVENTION: Structural Gene of Membrane-Bound
TITLE OF INVENTION: Alcohol Dehydrogenase Complex, Plasmid
TITLE OF INVENTION: Containing the Same And Transformed Acetic Acid
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fishaut, Holtz, Goodman & Woodward, P.C.
STREET: 600 Third Avenue
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10016-2088

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 inch, 0.72 mb
COMPUTER: IBM PC compatible (NEC PC-9801 ES)
OPERATING SYSTEM: MS DOS
SOFTWARE: ASCII Form
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/985,458
FILING DATE: 19921203
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/658,221
FILING DATE: 20-FEB-1991
APPLICATION NUMBER: 73440/1990
FILING DATE: 26-MAR-1990
ATTORNEY/AGENT INFORMATION:
NAME: Goodman, Herbert
REGISTRATION NUMBER: 17081
REFERENCE/DOCKET NUMBER: 910134/HG
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)972-1400
TELEFAX: (212)370-1622
TELEX: 236268

INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 738 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: MATURE PEPTIDE
LOCATION: 36 to 738
IDENTIFICATION METHOD: N-terminal sequences of the
IDENTIFICATION METHOD: purified protein having a molecular weight of about

```

IDENTIFICATION METHOD: 72,000
ORIGINAL SOURCE:
ORGANISM: Acetobacter alioacetigenes
STRAIN: MH-24
PUBLICATION INFORMATION:
AUTHORS: Tamaki, Toshimi;
AUTHORS: Fukaya, Masahiro;
AUTHORS: Takemura, Hiroshi;
AUTHORS: Takemura, Kenji;
AUTHORS: Okumura, Hajime;
AUTHORS: Kawamura, Yoshiya;
AUTHORS: Nishiyama, Makoto;
AUTHORS: Horinouchi, Suenaru and
AUTHORS: Beppu, Teruhiko
TITLE: Cloning and Sequencing of the Gene Cluster
TITLE: Encoding Two Subunits of Membrane-Bound
TITLE: Alcohol Dehydrogenase from Acetobacter
JOURNAL: Biochimica et Biophysica Acta.
VOLUME: 1088
PAGES: 292-300
DATE: 1991
US-07-985-458-3

```

Query Match 17.7% Score 479; DB 1; Length 738;
Best Local Similarity 25.9% Pred. No. 2e-37;
Matches 153; Conservative 77; Mismatches 230; Indels 130; Gaps 16;

```

QY 10 SAGLLALLAAPFAQVPTVDE-----LANPAGEWISYGOHONYRSPLOIT 61
Db 17 TACTICALISGATWASADGQATGEALIHDDHGNMTYGRYSQORSPIDQNR 76
QY 62 ENVGQLQVWARGMOPGRVQY-TPLIHGVYLANPBDVYQAI DATGGLIHERROL 119
Db 77 SNVGNKLAWYDLDTNNGEGTEPLVDGVYATYTWKSMKAVDATGRLMSYDPRV 136
QY 120 NIA-----TLNS-----FGI----- 129
Db 137 NIADGCCDYVRGAAVNGKVFYFGFDRLIALDAKTKLVMSVNTIPPEAEKORSY 136
QY 130 -----VANGVIVAGSTCOYSPFGC--FVSGHDSATGELMRNFIRAGEBD 175
Db 197 TVDGAPRIAKGRVILGN--GGSEFGARGVSAFDETGKVDREFTYVNPKNPEPDA 254
QY 176 -----ETGNDVEARWMTGA-----WQITVDPVYTNLVHGYSTAVGASEP 218
Db 255 VLMNKAYQWMS-----PTGAWTRQGGGYWDSIVDPVADLVLYGVNGSPWYKTR 307
QY 219 GTPCGTLYGNTFRFAVRPDTGETVWRHQTLPDMMWDOCEFTFEMWTVVAVQSP 278
Db 308 SEKGDNLFSGIYALKPRTGEYVHHPQETIMDQWDTISDQITLTLPT----- 357
QY 279 SINPNATGERRYLVGPCKTGTMOFDETEGFLMARDTNYONMIESID-ENGITV 337
Db 358 -----NGETRYVYHARKNGEYFIIDAKTGEFISKNVYVMAASGLDPTGRPI 410
QY 338 DALKELDEYVVCPTFLGGRDMPASALNPDGSIYFIPLNVCYDMAAADOQET-S 396
Db 411 DALYTLTGKREWGIPDGLGHNFAAASPOKTLVYIPAQOVPPVYDPOKGRKHH 478
QY 397 NTSN---NVTKLPKGRDMIGRIDDAIDISTGRILMSVEEAANYSPLVSTGGV 452
Db 479 DDNDPOHKADKQOFLDKLGVAMPDPOKQOATFVDBKGFWMGGLATAGVLFQGL 538
QY 446 LFNGGTDYRFRALSQETGETIMOTRLATYASGOAISYEVDGMOYVAIAGG 495
Db 531 LFOGLANGEFHAYDATNGSGLFFHADSGIAPVYTYLANGKQYVAEYVG 580

```

RESULT 3
US-09-296-284-25
Sequence 25, Application US/09296284A

```

1  APPLICANT: TOMIYAMA, No. 61271561bunt
2
3  TITLE OF INVENTION: D-Sorbitol Dehydrogenase Gene
4  FILE REFERENCE: D-Sorbitol Dehydrogenase Gene
5  CURRENT APPLICATION NUMBER: US/09/136,251A
6  CURRENT FILING DATE: 1998-08-19
7  EARLIER APPLICATION NUMBER: EP 97114432.4
8  EARLIER FILING DATE: 1997-08-21
9  NUMBER OF SEQ ID NOS: 9
10 SOFTWARE: Patentlin Ver. 2.1
11
12 SEQ ID NO 2
13
14 LENGTH: 739
15
16 TYPE: PRT
17
18 ORGANISM: Gluconobacter suboxydans
19
20 FEATURE:
21 NAME/KEY: SIGNAL
22 LOCATION: (1)..(24)
23
24 US-09-136-251-2

```

Query Match	11.8%	Score 319;	DB 3;	Length 739;
Best Local Similarity	21.8%;	Pred. No. 5,5e-22;		
Matches 149;	Conservative	197;	Indels 266;	Gaps 32

Dd	86	SOVPAMAPQOSANRQKDWAIYVRDHOQRYSPLSEITPEENASKLVAAYHMGSLTNRG	127
Oy	79	KV-----OVTEPL-IHDSVMTLANPGVIOAIDAKTGDILIMEHR-----ROLPIATLNSF	204
Dd	146	OYNKMAAETPTIKGDDLEYCSAMNDLIK-LDPATGKQIMRNVDVKHSLPTAACRGV	204
Oy	128	GIYANGVIVAGSTQOYSPPFGCVSGH-----DSANGELMNRHYEIPRAGE-----	172
Dd	205	TYETSSVYBEGQPCHE-----NRLIEGTLMDLMAVDDETD-----FCPNHGCGGVNLM	254
Oy	173	-----EGDETW-----GNDYEAKMTGAW-----	191
Dd	255	QGLGESVPGCFVSKTAPRPVINGVYVNHVELDQCRMAPSGVIKRYDASGKRYWANDVN	314
Oy	192	-----GOITYDPYTNLVHGXAVRPASTORCPFGGTLGVTNR-----FAVRPD	237
Dd	315	NSGRSOPAY-RVTVIYAVERRIPGLPDRORQG-----GSRLRPDRNSADYSAKLSKD	366
Oy	238	-----TGLTWHRHQTLPRDNMDQECTEEMVYTVWDVQSTEGEGSLNP	282
Dd	367	AENKVSAAVAIDVKTSERWVFOFANHDDVMDYDIGSOATL-----MDMPG-----P	413
Oy	283	NAATGERRVLTVGPKCTGIMMOQFDEATEFIMARDJNYNMIESIDEN-----GIY-----	333
Dd	414	DGQTVPALIM-----PYRKQOTFVILDRTRTG-----PLUVEERRAPSPGVLPDPR	459
Oy	334	-----TVNEDAI-----IKELD-----VEY	348
Dd	460	RSPTOPMVSVPKALRYPDLEKTMGMSPIDLFCRIKFRANVYGEFTPPSYDKPWIEY	519
Oy	349	DVCPTEFLGGRMPSAALNPDSGIYFTPLN-NVNCYDMAAVDELFMS-----	393
Dd	520	-----PCYNGSDWMSYSDQSGILIANWNTTPYDQULYTRKABSLGIAMPIDPBNRPGG	576
Oy	394	-----DYVNTSNVTKLPECKDMIGRIDAIDISIG-RILMS-----	427
Dd	577	GGAEGNGAMDGPICIVYTPPMQDY-TGMHCNPP-----YGMITVLDIMRHOOKYLMHPL	631
Oy	428	-----VERAANSPYLSTGGGVLENG-GTDRYRALSOETGETLMQTR	470
Dd	632	GSTARANGPWGLPTGLPWEITGTPNNGSVYTGGLIFITGAATDNQIRALDEHTGKVWSAV	691
Oy	471	LATVASGOAISYEVDGMOYVAIAGG	495
Dd	692	LPFGGQGANPMTYEANGHOYVAIAGG	716

RESULT

```

US-08-868-786-2
; Sequence 2, Application US/08868786
; Patent No. 5998701
GENERAL INFORMATION:
APPLICANT: Kawchuk, Lawrence M.
APPLICANT: Armstrong, John
APPLICANT: Lynch, Dermot
APPLICANT: Knowles, Richard
TITLE OF INVENTION: Potatoes Having Improved Quality
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSER: Greenlee, Winner and Sullivan, P.C.
STREET: 5730 Manhattan Circle, Suite 201
CITY: Boulder
STATE: Colorado
COUNTRY: USA
ZIP: 80803
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/868,786
FILING DATE: 04-JUN-1997
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/036,946
FILING DATE: 10-FEB-1997
ATTORNEY/AGENT INFORMATION:
NAME: Sullivan, Sally A
REGISTRATION NUMBER: 32,064
REFERENCE/DOCKET NUMBER: 8-97
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 499-8080
TELEFAX: (303) 498-8089
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 966 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-868-786-2

Query Match
Best Local Similarity 3.9%; Score 106; DB 2; Length 966;
Matches 87; Conservative 56; Mismatches 154; Indels 156; Gaps 19;

QY 25 VPPVDELIANPAGWISGQNGENRHSPLQITTEVNGQLV-----MARGM 75
Db 613 VPRRWIRFCNPLSATITWTGTEDW-----VKTETLAELOKFNEDNEDLQNEWRK 666
QY 76 QPGVQVYPLIHGCVMIANPAGVIAIDAKTDILWEHRRLPVIATLNSFGIVANGVI 135
Db 667 RSKIKIVASFLEKTKGTGYVP--DAMPDIQVR--IHEKTKQL-----LNFGLV----- 712
QY 136 VAGSTCOYSPFGCFVSGHSDATGEELMRVYFLPAGBEGDETWGNDYEA---RMTGAM 191
Db 713 -----YRYKKM-----KEMTAERKTN-FVPRVCIJFGKAFATYVQAIRYKFTITDVG 759
QY 192 GQITVD-----PVTNLIYHGSAAVGPASEQR---GTPG--- 222
Db 760 ATINHPPEIGDLKVFVVDYVNSVAVELLIPASDLSHISTAGMEASGSNNKFMANGCI 819
QY 223 --GLTGTWTRPAVRPDGELVWHRQLLPDRNDQECTFEEMVTNVVQSTEMEGLQSI 280
Db 820 QIGTLDGANVE--IREEGE----- 837
QY 281 NNNAATGERVLTGVPCKTGTMMQFDAETGEFLMARDTNVYNNMISIDENGALVYNEDAI 340
Db 838 -----ENFELFGAQAHEIAGLKRERADGFV--PDEFEEVKEFVNSGAFGSYNDVL 888

```

```

QY 341 LKELD-----VEYDVCPTFIGRDMPSAALNDSDIYFIPLNNVCYDMMAVDOEFTSM 393
Db 889 IGSLENGEGRGRADY-----FLVGRDPS-----YIEQGEVDEAYDQKRWITM 933
QY 394 DYVNTSNVTKLPKGRDM-----IGRIDAIDIS 420
Db 934 SILNTAGSYKFSDDRTIHEYAKDWMNEAVEIA 966

RESULT
US-08-219-262B-10
; Sequence 10 Application US/08219262B
; Patent No. 5788970
GENERAL INFORMATION:
APPLICANT: VAKHARIA, VIKRAM
APPLICANT: SNYDER, DAVID B
APPLICANT: MENDEL-WHERSAT, STEPHANIE A
TITLE OF INVENTION: CHIMERIC INFECTIOUS BURSAL DISEASE VIRUS
TITLE OF INVENTION: CDNA CLONES, EXPRESSION PRODUCTS AND VACCINES BASED
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSER: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT
STREET: 1755 S. JEFFERSON DAVID HIGHWAY, FOURTH FLOOR
CITY: ARLINGTON
STATE: VIRGINIA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/219,262B
FILING DATE: 29-MAR-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NORMAN F
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 2747-047-27
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 413-3000
TELEFAX: (703) 413-2220
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 1012 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: protein
ORGANISM SOURCE: protein
ORGANISM: Infectious bursal disease virus
STRAIN: OH
US-08-219-262B-10

Query Match
Best Local Similarity 3.9%; Score 105; DB 1; Length 1012;
Matches 84; Conservative 45; Mismatches 130; Indels 120; Gaps 19;

QY 199 VTNLYVSGTAV-----GPAS-----ETORGPGTGLVGTNTRF 232
Db 1 MTNLDHQIQLVPIRSLMLPTGPASIDPDILKHTLRSETSTYNTLVGDTGSLIVFF 60
QY 233 AVRPDTGELVWHRQTLPRD---NMDOECTFEEMVT---NVVQSTEME 273
Db 61 PGFP--GSVGAHTTLOSNGSYQFDQ---MLTPAQLMPSVYNCRLVRSLSLTVRSSTL 113
QY 274 MEGLOSINP--NNAATGERVLTGVPCKTGTMMQFDAETGEFLMARDTNVYNNMISTE----- 327

```

Fri May 24 11:27:29 2002

walick-934-128.pep.rai

Page 5

Db 114 PGVYALNGTINAVT-----FGSLSEL---TDYSYNGMSATNIN 152
OY 328 DENGIVTNEDELLKELDEYDVCPTFLGRDMPKALNPDGSIYFPLNNVC-----380
153 DKIGNVLVEGYVTLSTPLSYDLSTYRLGD-PIPAAGLDP-----KLMATCDSDRPR 204
OY 381 -YDMAVADQ-EF-----TSMDEVNTSNVTKLPKPKDM-GRIDAIDISGRTLM-426
Db 205 VYVTADAEYQFSSOLIPSGVKTILFTANIDALTSVGGELIFSOVTHSLEDVTTIF 264
OY 427 -----SVERAANYSPVLSTGGVLFN--GGDRYFRALSOETGETLMOTRLATYAS 476
Db 265 IGFDTETVYKAVATDFGLTTGTNNLVPNLGPTSETIQPTSMKLEVYTKRGGT--A 322
OY 477 GOAISYEVDGMOYVAIAGG 495
Db 323 GDPISMTVSGTLAVTIYVG 341

RESULT 7
US-09-031-655-10
Sequence 10, Application US/09031655
Patent No. 6017759

GENERAL INFORMATION:
APPLICANT: VAKHARIA, VIKRAM

APPLICANT: SNIDER, DAVID B

APPLICANT: MENGEL-WHERSAT, STEPHANIE A

TITLE OF INVENTION: CHIMERIC INFECTIOUS BURSAL DISEASE VIRUS

TITLE OF INVENTION: CDNA CLONES, EXPRESSION PRODUCTS AND VACCINES BASED

TITLE OF INVENTION: THEREON

NUMBER OF SEQUENCES: 15

CORRESPONDENCE ADDRESS:

ADDRESS: 1755 S. JEFFERSON DAVID HIGHWAY, FOURTH FLOOR

CITY: ARLINGTON

STATE: VIRGINIA

COUNTRY: USA

ZIP: 22202

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/031,655

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA: US 08/219,262

APPLICATION NUMBER: 29-MAR-1994

FILING DATE: 29-MAR-1994

ATTORNEY/AGENT INFORMATION:

NAME: OBLON, NORMAN F

REGISTRATION NUMBER: 24,618

REFERENCE/DOCKET NUMBER: 2747-047-27

TELEPHONE: (703) 413-3000

TELEFAX: (703) 413-2220

TELEX: 248855 OPAT UR

INFORMATION FOR SEQ ID NO: 10:

SEQUENCE CHARACTERISTICS:

LENGTH: 1012 amino acids

Matches 84; Conservative 45; Mismatches 130; Indels 120; Gaps 19;
OY 199 VTNLVHGSTAV-----GPAS-----ETORGPGGTLYGINTRF 232
Db 1 MTNLMQHTQOIIYPIFRSLMPTGTGASIPDDTLKHTLRSESTYLVIGDYGSLYVF 60
OY 233 AVRPDGEIYWRHQTLPD-----NMDOCTEFEMAVT-----WVDQPESTE 273
Db 61 PGFP--GSVGAHYTLSSNGSOPFO-----MLTLAONLPVSYNYCRLYSRSLTYRASSYL 113
OY 274 MEGSOSIMP--NATGERHYLGVCCKTGTMMQPDATGEFIMARDTNQNMTEST-----327
Db 114 PGVYALNGTINAVT-----FGSLSEL---TDYSYNGMSATNIN 152
OY 328 DENGIVTNEDELLKELDEYDVCPTFLGRDMPKALNPDGSIYFPLNNVC-----380
153 DKIGNVLVEGYVTLSTPLSYDLSTYRLGD-PIPAAGLDP-----KLMATCDSDRPR 204
OY 381 -YDMAVADQ-EF-----TSMDEVNTSNVTKLPKPKDM-GRIDAIDISGRTLM-426
Db 205 VYVTADAEYQFSSOLIPSGVKTILFTANIDALTSVGGELIFSOVTHSLEDVTTIF 264
OY 427 -----SVERAANYSPVLSTGGVLFN--GGDRYFRALSOETGETLMOTRLATYAS 476
Db 265 IGFDTETVYKAVATDFGLTTGTNNLVPNLGPTSETIQPTSMKLEVYTKRGGT--A 322
OY 477 GOAISYEVDGMOYVAIAGG 495
Db 323 GDPISMTVSGTLAVTIYVG 341

RESULT 8
US-08-868-786-6
Sequence 6, Application US/08868786
Patent No. 5998701

GENERAL INFORMATION:
APPLICANT: Kawchuk, Lawrence M.

APPLICANT: Armstrong, John

APPLICANT: Lynch, Dermot

TITLE OF INVENTION: Knowles, Richard

TITLE OF INVENTION: Polatoes Having Improved Quality

TITLE OF INVENTION: Characteristics and Methods for their Production

NUMBER OF SEQUENCES: 10

CORRESPONDENCE ADDRESS:

ADDRESS: 5730 Manhattan Circle, Suite 201

CITY: Boulder

STATE: Colorado

COUNTRY: USA

ZIP: 80803

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/868,786

FILING DATE: 04-JUN-1997

CLASSIFICATION: 800

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 60/036,946

FILING DATE: 10-FEB-1997

ATTORNEY/AGENT INFORMATION:

NAME: Sullivan, Sally A

REGISTRATION NUMBER: 32,064

REFERENCE/DOCKET NUMBER: 8-97

TELEPHONE: (303) 499-8080

TELEFAX: (303) 499-8089

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 974 amino acids

TYPE: amino acid

Query Match 3.9%; Score 105; DB 3; Length 1012;
Best Local Similarity 22.2%; Pred. No. 0.42;

TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-868-786-6

Query Match
Best Local Similarity 3.8%; Score 103.5; DB 2; Length 974;
Matches 89; Conservative 65; Mismatches 166; Indels 161; Gaps 19;

14 LALLAAPAAQVPTVDLLANPPAGE---WISTQONEN-----YRHSPLTOIT 60
Db 579 LCVSGHANGVAEIHSEIVADEVFNEFKLWPEKQNKNTNVTFRMLSCNELSEIT 638
QY 61 TENVG-----OLQVYAKQOPKYOYVPLHGDVMTLANGDV 99
Db 639 TKWTGSDMLVNTKELKFKADNEELQSEWRKAKGNKKIYSLIKETGYVSP-DA 697
QY 100 IQAIDAKTGLIMEHRQPLNATLNSFGIVANGVIYAGSTCOYSPGCEVSGHDSATGE 159
Db 698 MEDVOIKR---IHEYKROL-----INFGIV-----YRKKM-----KEMSP 731
QY 160 ELARNYFIPRAGEGDETWGNDYEA---RMKGAMQOITYDPV-----YRHSPLTOIT 60
Db 732 EERKEKFPVPCIFGSKANATYVAKKRIKFTIDYGETVNHDEIGDLLKVFVVDIYVS 791
QY 200 -----TNLVHGSTAVGPASETOR---GTPG---GLYGNTRAVRPDGETIYWR 244
Db 792 VAEVLIPGSELQSHISTAGMEASGTSNMKFSMNGCLITIGDGAWE---IREEVEG--- 845
QY 245 HQTLPRDNDOCTEFEMVATVNDVQSTEMEGLOINPAATGERVLTGVCCKTGMQ 304
Db 846 -----DNFFLEGAQAHIEIAGLKRERAEKFPVDPDFEEVKAFITGV---FGTY-- 891
QY 305 FEAETGEFLMARDNTYONMIESIDENGCIYVNEAIIKELDEYDVOPTLGRGMDPSAA 364
Db 892 -----NYEELGSLGEGN-----EGGRADYFLVCKDPP-- 920
QY 365 LNPDSGIYFPLNNVCYDMAVDOFTSMVYNTSNVTKLPKPGDM-----IGRIDALD 418
Db 921 -----YIECODKVDAYVRQKKTKMSTLNTAGSFKFSORTIHOYARDIMRIEPE 972
QY 419 I 419
Db 973 L 973

RESULT 9
US-08-928-361B-11
Sequence 11, Application US/08928361B
Patent No. 6071518
GENERAL INFORMATION:
APPLICANT: Petersen, Carolyn
TITLE OF INVENTION: PEPTIDES, POLYPEPTIDES, GLYCOPROTEINS,
TITLE OF INVENTION: THEIR FUNCTIONAL MUTANTS, VARIANTS, ANALOGS AND FRAGMENTS
TITLE OF INVENTION: FOR TREATMENT AND DETECTION/DIAGNOSIS OF CRYPTOSPORIDIUM
NUMBER OF SEQUENCES: 30
CORRESPONDENCE ADDRESS:
ADDRESSEE: PETERS, VERNY, JONES & BIKSA
STREET: 385 Sherman Avenue, Suite 6
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306-1840
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/928,361B
FILING DATE: 12-SEP-1997
CLASSIFICATION:

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/026,062
FILING DATE: 13-SEP-1996
ATTORNEY/AGENT INFORMATION:
NAME: VERNY, HANA
REGISTRATION NUMBER: 30,518
REFERENCE/DOCKET NUMBER: 480,76-1(HV)
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-324-1677
TELEFAX: 650-324-1678
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 1042 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-928-361B-11

Query Match
Best Local Similarity 3.8%; Score 101.5; DB 3; Length 1042;
Matches 138; Conservative 71; Mismatches 222; Indels 229; Gaps 36;

25 VTPVDELLANPPAGEMISTYQONEN-----YRHSPLTOITENYQOLVWARGMOP 77
Db 301 IDPTTG-LFPNPPGHLINPTNNNTMDSFAGAYKTAVANGIKITDVTGLPVDITGLPK 359
QY 78 GRQVQVPLIHGDVMTLANPGDVIAIDAKTGLIMH-----RROLPNATL----- 125
Db 360 DRYSDIP-----FNSYTGEL---VDPTGKPINNYAGIYSGKRGILPDEDNGMLD 409
QY 126 ---SEGIYANGVI-----VAGST---COYSPGCEVSGHDSATGEELARNYFIPRAG 171
Db 410 PSTKLPIDGNNQVLPNPEPNTSTYSGSTGSKRP-GIPVNGGVPVDE-----AK 459
QY 172 EGGDETMNGNIEARWNTGAMGOITTDVPTNLVHYGSTA---VGRASEYQKGTG----- 222
Db 460 DQAK--GKD---GLIVPPTNSINKDPVTNQTNTGNTGNIINP---ETGKVIYPSLPSLSLN 512
QY 223 -----GLYGTNTFAVRPDGETIYWRQTLPRDN---MDQCTFEMVNTN 265
Db 513 YPSEFNTPOQTDITGKPDVDTGTPYDPSGTEIIDATYKLPVGSVAGDELTEVNTIT 572
QY 266 -----VDVQ---PSTMEGLQS-----INP-----NAAG-----ERRV 291
Db 573 DEVTGLPIDLETGLPRDPSGLPOLPNGTLVPSNKKPIPSHSGFINGISGEQSHKDP 632
QY 292 LTGVPCCKTGTMQPD-----AETGEFLMARDNTYONMI---ESIDENGIYVNEADIL 341
Db 633 STGKFLDNTGLPDEDGSLINPETGDKLQSGHSGTTPMPVPGKPOENCGIMTPPO-IL 691
QY 342 KELDY---EYDQCP-----TFE---GGRWMPSEAL----- 365
Db 692 EALKLFTSNEVNISPRPSDAVPPDRPTNTWNNKISGQTFQVQDKTIGSASAVIHAL 751
QY 366 -----NDQSGIYFPLNNVCYDMAVDO-----EFTSMVYNTSNVTKLP 406
Db 752 GTPYQIDPTTGLPSDPSGLPIPGFNVLVDPQNGEDIKSVYVSLVYKKNITTEAY 811
QY 407 GKDMIGRIDALISTGRILMSVERAANYSPIVSTG-----GVLFNGGIDRYFRA--- 457
Db 812 GL-----PVPKGTGPIIDI---SYLPAKNGELIDPISGKYFSIAGTISKAG 859
QY 458 ---LSOETGETIMQT-----RLATVASGOAISEYVQGMQYVIAAG-----GVV 497
Db 860 SQSKSDSGNPIIDPSNMPYDPKTKLIDPESGLAIDNSVSGV-FATVPGTAPKPKGV 918

RESULT 10
US-08-928-361B-5
Sequence 5, Application US/08928361B
Patent No. 6071518

GENERAL INFORMATION:

APPLICANT: Petersen, Carolyn
 TITLE OF INVENTION: PEPTIDES, POLYPEPTIDES, GLYCOPROTEINS,
 THEIR FUNCTIONAL MUTANTS, VARIANTS, ANALOGS AND FRAGMENTS
 TITLE OF INVENTION: FOR TREATMENT AND DETECTION/DIAGNOSIS OF CRYPTOSPORIDIUM
 TITLE OF INVENTION: SPECIES INFECTIONS
 NUMBER OF INVENTIONS: 30
 CORRESPONDENCE ADDRESS:
 ADDRESS: PETERS, CARYN, JONES & BIKSA
 STREET: 385 Sherman Avenue, Suite 6
 CITY: Palo Alto
 STATE: CA
 COUNTRY: USA
 ZIP: 94306-1840
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/928,361B
 FILING DATE: 12-SEP-1997
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 60/026,062
 FILING DATE: 13-SEP-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: Verry, Hana
 REGISTRATION NUMBER: 30,518
 REFERENCE/DOCKET NUMBER: 480.76-1(HV)
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 650-324-1677
 TELEFAX: 650-324-1678
 INFORMATION FOR SEQ ID NO: 5:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1837 amino acids
 TYPE: amino acid
 STRANDEDNESS:
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-928-361B-5

Query Match 3.8%; Score 101.5; DB 3; Length 1837;
 Best Local Similarity 20.9%; Pred. No. 2.4; Indels 229; Gaps 36;
 Matches 138; Conservative 71; Mismatches 222

25 VPEVTELLANPPAGEMISYQNOE-----NYRHSPLQITTEVNGQOLVWAKMOP 77
 1096 IDPTTG-LPPNPTGHLINPTNNMTDSSFGAYKYAVSNGIKTDNYGGLPVEITGLPK 1154
 78 GKQVPELIDHGVYIANGDYIAIDAKTGDLIMEH-----RROLPNATLN----- 125
 1155 DPVSDIP-----FNSITGEL---VDPESTGKPINNYAGIYSGKRLPIEDENGWLF 1204
 126 ---SFGIVANGV-----VAGST---CQYSPGCEVSGHDSATGEELMRNTFIRAG 171
 1205 PSTKLPIIDGNOLVNPETNSYSGSTSKPKP-GIPYNGGVYDPDE-----AK 1254
 172 EESDEIWNDEYKARMTGANGQITTYDPTNLYHGSTA---VGPASERGRTPG----- 222
 1255 DQADK---GKD---GLIYPTNSINKDPTNTQYISNTGNINP---ETKAVIPGSLPGSLN 1307
 223 ---GTLGTNTRFAVAPRPTGELIWMRHQTLPRDN---MDOCTEEMAYTN 265
 1308 YPSENTPQDDEITGKPYDVTGTPDPTGELIDPATKLPISGAVAGDEILTEVLTNTT 1367
 266 ---VDVQ---PSTEMEGLS-----INP-----NATG---ERRV 291
 1368 DEVTGLPIDELTGPRDPVSGLPQLPNGTLVDSNKKPIPGSHSGFINGTSESGHEKDP 1427
 292 LFGVCKTGTYMOPD-----AETGEFLMADTYNOMNI---ESIDENGIVTVNEDAIL 341

1428 STGKPLDPNTGLPEDEDSGLINPETGDKLQSGHSTFMPYKRGKENGIMTPEO-IL 1486
 342 KEDV-----EYVCP-----TFL---GGRDWPSAAL----- 365
 1487 EALNKLEPTESEVNSIPRPSDAPDRPTWNNKISGOTFOVDGKKTIPGSAASVHTAL 1546
 366 ---NPDSGIYFIPLNVCYDMMAVDQ-----ETSMYVYNTSVTKLP 406
 1547 GTPQOTDPTTGLSDSPETGLPIPGFNVLVDPOTEQIKGSVYVSLYKREKNTVEANAY 1606
 407 GKDMIGRIDAIDISTGRTLMSVERAANSPVLSTG-----GGVLFNGIDRYEFA--- 457
 1607 GL-----PYDPKTPGPIDPI-----STLPFAKNGBELDIPSGKFFSGSIAGFISGRAG 1654
 458 ---LSOETGETLMOT-----RLATVASGAQISYEVDMQOTVAILG-----GGV 497
 1655 SOSKSDSEGNPDPSTNMPYDPRTGKLDIPESGIAIDNSVSGV-FATVGTAPDKKGV 1713

RESULT 11
 US-09-540-245A-16
 Sequence 16, Application US/09540245A
 Patent No. 6270984

GENERAL INFORMATION:
 APPLICANT: Goodman, Corey
 APPLICANT: Kid, Thomas
 APPLICANT: Brose, Katja
 APPLICANT: Tessier-Lavigne, Marc
 TITLE OF INVENTION: Modulating Robo: Ligand Interactions
 FILE REFERENCE: 898-031-3
 CURRENT APPLICATION NUMBER: US/09/540,245A
 CURRENT FILING DATE: 2000-03-31
 PRIOR APPLICATION NUMBER: 60/065,544
 PRIOR FILING DATE: 1997-11-14
 PRIOR APPLICATION NUMBER: 60/081,057
 PRIOR FILING DATE: 1998-04-07
 NUMBER OF SEQ ID NOS: 20
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 16
 LENGTH: 1381
 TYPE: PRT
 ORGANISM: Drosophila melanogaster
 US-09-540-245A-16

Query Match 3.7%; Score 99.5; DB 4; Length 1381;
 Best Local Similarity 20.6%; Pred. No. 2.4; Indels 125; Gaps 20;
 Matches 90; Conservative 51; Mismatches 170

97 GDVIO-----AIDAKTGDLIME-----HROLPNATLNSFGIYANGVIYAGS 139
 626 GDVIELNASVYDSTSMKLTWQIINGKYVEGFYVAROLPNPIVNNPAPVSNTPILGS 685
 140 TCQYSPGCEVSGHDSATGEELMRNT-FIPRAGEEDETMGN-----DYEARMT-- 188
 686 TS-----TSASASASASALISTKPNIAAKRGELINOSGGAPPLMKRYMLTIL 737
 169 ---GANGQITTYDPTNLYH-----GSTAVGPASEYORGRPGTGLGTMR 231
 738 NCGGASCT---ITGIVQYTYLFEFFIYFVKYSGKPKSRIRARTLEDVPSRAPYGEA- 793
 232 FAVRPDTEIYWRHQTLPDWMDOCTEFEMMYTNVDOPSTEMEGLSINPNAATGERRV 291
 794 ILLNSSAVFLWKAPEL-KDRHGYLANIHYIVKID-----TANFSRI 836
 292 LFGVCKTGTYMOPDAETGEFLMADTN---YONMIESIDENGI-----VYNEDAILK 342
 837 LINV-----TIDAASPTLVLANITGEGVATVGAAGNNAAGVCPATLIDPLTK 888
 343 ELDEVYDCPTFLGGRD-----WPSAALNPDGSIYFIPLNVCY---DMAAVDOE 389
 889 RLD-----PFINORDHVNDVLTQPFITLLGAILAVLMISFGANVFYRKRMHMKQSA 941

QY 390 FTSMDVYNTSVYKLPCKDMIGRIDALDISTGRTLM-----SVERA---AANYSPVL 439
 DB 942 LMTMGNHTSDVLMKPSLSARNGCYWLDSTGGMWRPSPGGSDLEMKDHIADYAPVC 1001
 QY 440 ----STGGVLENGCT 451
 DB 1002 GAPGSPAGGTSSGGS 1017

RESULT 12
 US-07-642-734C-4
 Sequence 4, Application US/07642734C
 Patent No. 5824513

GENERAL INFORMATION:
 APPLICANT: Katz, L
 APPLICANT: Donadio, S
 APPLICANT: Mcalpine, J B
 TITLE OF INVENTION: Recombinant DNA Method for Producing
 Erythromycin Analogs
 NUMBER OF SEQUENCES: 27
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Edward H. Gorman
 STREET: Abbott Laboratories D377/AP6D-2 One Abbott
 STREET: Park Rd
 CITY: Abbott Park
 STATE: IL
 COUNTRY: US
 ZIP: 60064-3500

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/07/642,734C
 FILING DATE: 17-JAN-91
 CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
 NAME: Dancigers, Andreas M
 REGISTRATION NUMBER: 32652
 REFERENCE/DOCKET NUMBER: 4952.US.01
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 708-937-9396
 TELEFAX: 708-938-2623

INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 3567 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-07-642-734C-4

Query Match 3.7%; Score 99.5; DB 2; Length 3567;
 Best Local Similarity 22.9%; Pred. No. 11;
 Matches 125; Conservative 48; Mismatches 221; Indels 153; Gaps 26;

QY 21 AFAQVTPVDELLANPPAGEW-----ISYGNQENYRHSPLTQITTEVNGQLQVW 71
 DB 2911 AFAPIAVTDBRLARVPDG-WSDADAANVPLAV-----TTHAYLHDLAIGRAQSVLIH 2964
 QY 72 ARGMQPGKQVYPLIHDSQMYLANPGDYIOAIDAKTGLIWEHRRQI-----PNIATLNSF 127
 DB 2965 AAGGCGMAAVALARRAGAEVLATAG-----PAKHGTL-----RALGLDEHIASGRET 3013
 QY 128 G-----IVAN--GVIVAGSTQYSPFGCFYS--CHDSATGEELMRNFIP 168
 DB 3014 GFARKFRERTGGRGVYVINSITGLDESDADLLAEDGVEVEMKTDLDRAQDGRGRYAP 3073
 QY 169 -RAGEGSDTGWNDY--EARMATGANGQITTYDPTNIV-----HYGSTAV- 210
 DB 3074 FDLGEGDRLGEILREVYGLGA-GELDLRPVSAMELGSAPAAALQMSGRHVGKLVLT 3132

QY 211 GPASRTQRT---PGTLTGNTFRPAVPDTGELVWRHQT-PRDMNDQCTFEEMAVTVN 266
 DB 3133 QPAPVDDPDTGLITGTG--GTLGRLLARHLVTEHGVHLLVSRGADAGSDELAAELE 3190
 QY 267 DVQPSYTEMELQSIQNPNAF-----GERVLTNGVCKTGTWQDAETGEPLNARDNTION 322
 DB 3191 DLGASAEIAACDTADRDALSLALDGLRPLTGVAAGV----- 3230
 QY 323 MIESIDENGIVT-VNEDAIILELDEVYDQPTFLGGRDPSALNDPSGIYFPIPNVCY 381
 DB 3231 ----ADGLVTSIDEPAVEQVLRKVDAA-----NWLHET/ANTGLSFVLFSSNA 3276
 QY 382 DMKAVDQFTSMQDYNTSNV-----TKLPCK-----DMTGR 413
 DB 3277 SVLA-----GPGQGYAAANESLNALALRRTKRLPAKALGWLMAQASEMTSGLDRIAR 3332
 QY 414 IDAIDISTGRTLMSVERAANYSPVLSTGGVLENGCTDRYFRALSOETGETL-WQTRLA 472
 DB 3333 ----TGVAALPTERALLFDSALRRGGEVYFPLISIRSALEPVEYLKGMVRAK 3385
 QY 473 TVASGOA 479
 DB 3386 LRAAGQA 3392

RESULT 13
 US-08-439-009A-4
 Sequence 4, Application US/08439009A
 Patent No. 6004787

GENERAL INFORMATION:
 APPLICANT: Donadio, S
 APPLICANT: Katz, L
 APPLICANT: Mcalpine, J B
 TITLE OF INVENTION: Method of Directing Biosynthesis of
 Specific Polypeptides
 NUMBER OF SEQUENCES: 27
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Steven F. Weinstein
 STREET: Abbott Laboratories D377/AP6D-2 One Abbott
 STREET: Park Rd
 CITY: Abbott Park
 STATE: IL
 COUNTRY: US
 ZIP: 60064-3500

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/439,009A
 FILING DATE: 11-MAY-1995
 CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
 NAME: Casuto, Dianne
 REGISTRATION NUMBER: 40,943
 REFERENCE/DOCKET NUMBER: 4952.US.D1
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 847-938-3137
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 3567 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-439-009A-4

Query Match 3.7%; Score 99.5; DB 3; Length 3567;
 Best Local Similarity 22.9%; Pred. No. 11;
 Matches 125; Conservative 48; Mismatches 221; Indels 153; Gaps 26;

QY 21 AFAQVTPVDELLANPPAGEW-----ISYGNQENYRHSPLTQITTEVNGQLQVW 71

Fri May 24 11:27:29 2002

walick-934-128.pcp.ra1

Page 9

DB 2911 AFAPLIVDRLRLARVPDG-MSDADAAVPLAY-----TTAAVHLDLAGLAGAGSVLIH 2964
QY 72 ARGMPGKQVQPLIHGQVMTLANPGDVIAIDAKTGLDIMEHROL-----PNIATLNSF 127
DB 2965 AAGGVMAAVLARRAGAEVATAG-----PAKHGTL-----RALGDDDELASSRET 3013
QY 128 -----IVAN-----GVYAGSTCOYSPRGCEVS-GHSAATGEELMRYVFI 168
DB 3014 GFARKFRERTGSGVDVYVNSLTGELDESADLLAEDGVFEKGTDLRAGDFRGRYAP 3073
QY 169 -RAGEGDEFWGDY-FAARMGAMGOITVDPVNV-----HYGSTAV- 210
DB 3074 FDLGAGDRLGELRLVGVGLGA-GEEDRLVSAWELGSAAPALOHMSRGRHVKLVLT 3132
QY 211 GPASETORG-PPGTLVGTNTFRFVAPDPTGEIYWRHQT-PRDNWDOECTFEEMVTNV 266
DB 3133 OPAPVDPGTLITGT-CTIGRLRLARHLYEHGVRHLLVSRGADAPGSDRLAEIE 3190
QY 267 DVQSTEMEGLOSINPNAAT-----GERRVLTGVPCKTGTMMOFDAETGEFILMADTNYON 322
DB 3191 DIGASAEIACDTRADRLSALDGLRPLTGVVHAAGVL----- 3230
QY 323 MIESIDENGIVT-VNEDAILKELVEYDVCPTLGGDMPFSALNPDGCIYPIPLNNVCY 381
DB 3231 -----ADGLVTSIDEPAVEQVLRAKVDA-----WMLHLLFANTGLSFEVFLSSAA 3276
QY 382 DMAVDOETSMDYNTSNV-----TKLPPGK-----DMIGR 413
DB 3277 SVLA-----GGGQGVYAANESINLAALRRRGLPARALGMLMAQASEMTSGLSGRIRAR 3332
QY 414 IDAIDISTGRTMSVERAANSPVLSGGVYENGTDYFRALSOETGETL-WQTRLA 472
DB 3333 -----TGAALPTERLALFDSLRGGEVYFPLSINFSALRRAEVFEVLGMYRAK 3385
QY 473 TVASGOA 479
DB 3386 LRAAGOA 3392
RESULT 14
US-08-856-841-22
Sequence 22, Application US/08856841
Patent No. 6274307
GENERAL INFORMATION:
APPLICANT: ERWIN SOUTSCHEK
APPLICANT: MANFRED MOTZ
TITLE OF INVENTION: IMMUNOLOGICALLY ACTIVE PEPTIDES
TITLE OF INVENTION: OR POLYPEPTIDES FROM THE PARVOVIRUS B19
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROOKS HAIDT HAFNER & DELAHUNTY
STREET: 99 PARK AVENUE
CITY: NEW YORK
STATE: NY
COUNTRY: USA
ZIP: 10016
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" FLOPPY DISC
COMPUTER: AT&T - IBM COMPATIBLE
OPERATING SYSTEM: MS-DOS Version 6.2
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/856,841
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/214,658
FILING DATE: 16-MARCH-1994
APPLICATION NUMBER: US 07/917,096
FILING DATE: 4-AUGUST-1992
APPLICATION NUMBER: PCT/DE91/00106
FILING DATE: 8-FEBRUARY-1991

APPLICATION NUMBER: DE40038262
FILING DATE: 8-FEBRUARY-1990
ATTORNEY/AGENT INFORMATION:
NAME: ROBINSON, WILLIAM R.
REGISTRATION NUMBER: 27,224
REFERENCE/DOCKET NUMBER: LKR-9222-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 697-3355
TELEFAX: (212) 557-5635
TELEX: NONE
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 543
TYPE: AMINO ACID
TOPOLOGY: LINEAR
MOLECULE TYPE: PEPTIDE
DESCRIPTION: N/A
HYPOTHETICAL: N/A
ANTI-SENSE: N/A
FRAGMENT TYPE: INTERNAL
ORIGINAL SOURCE: SERUM FROM PATIENT WITH ACUTE
ORIGINAL SOURCE: INFECTION (ERYTHEMA INFECTIONUM)
IMMEDIATE SOURCE: GENETICALLY ENGINEERED PEPTIDE
POSITION IN GENOME: N/A
FEATURE:
NAME/KEY: N/A
LOCATION: N/A
IDENTIFICATION METHOD: amino acid analysis and
IDENTIFICATION METHOD: mass spectrometry
OTHER INFORMATION:
PUBLICATION INFORMATION:
AUTHORS: COSSART, Y.E.
AUTHORS: FIELD, A.M.
AUTHORS: CANT, B.
AUTHORS: WIDDOWS, D.
TITLE: PARVOVIRUS-LIKE PARTICLES IN HUMAN SERA
JOURNAL: LANCET
VOLUME: I
ISSUE:
PAGES: 72 - 73
DATE: 1975
DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO: 22:
US-08-856-841-22
Query Match 3.68; Score 97; DB 4; Length 543;
Best Local Similarity 21.0%; Pred. No. 0.9; Indels 146; Gaps 23;
Matches 96; Conservative 48; Mismatches 168
QY 41 WISYGNQENYRHSPLT-QITTEVNGQLVWARGMOPKQVQPLIHGVMYLA----- 94
DB 96 WRYLDENALNLFPSPLEPHQHLIENGSI-----APDALVY-----ISEIAVKQVT 141
QY 95 -NPGDVIOAIDAKTGLD-LTWEHROLP-----NIATL 124
DB 142 DKGGGVQVDTSTYGRLOMLVDHEKRYVLAGGGDTLAPELPIWYFPOQVAVLTGVQV 201
QY 125 NSRGIVANGIVAGSTCOYSPRGCEVSGHS-----ATGEELMRYVFI PRAGEEG- 174
DB 202 NTQGISGSKKLMSBSAF-----VYLEHSSPOLLGTGTSMSKXKPPYPRPELESCSO 256
QY 175 -----DETWGNDYEARMNTGAMGOITVD-----PVNULVYVYSTAVGPA 213
DB 257 HFYEMYNPLYSRLGVDPDLGDPKRSRLTHEDHAIOPNFMFGFLVNSV-----STKEGDS 313
QY 214 SETORGTP-GGTLVGT-NTRFAVRPDTGIVARHQTLPDRNMWDOECTFEEMVTVNDV-- 266
DB 314 SNTGAGKALTLGISTGSONTRISLRP--GPV-----SQYTHHMDT-----DKVYTGINKIS 362
QY 269 -----QPTSTEMEGLOSIN-----PNAAGERRVLTGVVCKTGT 301

```

Db 363 HQGTYGNAEDKEXQGVGRFPRNEKQLOGLNMHTYFPKKGTYQVTDQIERPLAVGS 422
QY 302 MMOPDAETGEF-IMAR-----DNYQNMIESIDENGIVAVNDAIILKELDE-IDVCEPTEL 355
Db 423 VMNRRAHYISQSLMSKIPMLDSEFKTFQFAALGCMGLHPPOPTLKOYAVGIMTYMTERK 482
QY 356 GGRWMSAALNPDGSIYFIP-----LNNVCYDMMAYD 387
Db 483 LGPRKATGRWNPQGVY--PRAACHLPYVLYXDPYAD 518

RESULT 15
US-08-700-651-5
; Sequence 5, Application US/08700651B
; Patent No. 6015882
; GENERAL INFORMATION:
; APPLICANT: PETERSEN, CAROLYN
; APPLICANT: LEBCH, JAMES
; APPLICANT: NELSON, RICHARD, C.
; APPLICANT: GUT, JIRI
; TITLE OF INVENTION: VACCINES, ANTIBODIES, PROTEINS, GLYCOPROTEINS, DNAS AND RNAS
; TITLE OF INVENTION: FOR PROPHYLAXIS AND TREATMENT OF Cryptosporidium parvum
; FILE REFERENCE: 480.19-4(HV)
; CURRENT APPLICATION NUMBER: US/08/700,651B
; EARLIER FILING DATE: 1997-08-14
; EARLIER APPLICATION NUMBER: 08/415,751
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 1721
; TYPE: PRT
; ORGANISM: Cryptosporidium parvum
US-08-700-651-5

```

Query Match 3 6%; Score 97; DB 3; Length 1721;
 Best Local Similarity 20.5%; Pred. No. 5.9;
 Matches 136; Conservative 78; Mismatches 215; Indels 234; Gaps 38;

```

QY 25 VTPVDELLANPAGENISYQNOE-----NYRHSPLQITTEBNVGOLOLVWARGMOP 77
Db 979 IDPTTG-LPENNPTGHLINPTNNMTDSSFAAYKAYVANGIKITDNV-----YGLPV 1029
QY 78 GKVOVTEPLIHGVMYLANPGVIAQDAKTDGLMEHRROLNIALINSFGIVA-----YGLPV 1029
Db 1030 GEITGLP-----KDGSDI--FPNSTGELVDPSTGKPINNSTA---GIYSGKRGIP 1076
QY 132 -----NGVI-----VAGSTCOYSP-FGCFVSGHDSAT-----GELMRNFTIP--RA 170
Db 1077 PIDEENGNI.FDPSYMLPIDGNMOLVNPETNSTVSGSTGTTKPKPGIPIVNGGGVVPDEBA 1136
QY 171 GEGDETWGNDYEAARMWTGANGOLTYDVPVNLVHYGSTA--VGPASETORGTGPG----- 222
Db 1137 KQOAKR--GKD-----GLIYPPINSINKDPVNTQYISMTGNIINP--ETGKVIPGSLPGSL 1189
QY 223 -----GLIYGINTRFAVVRPDTGEIWRHOTLPRDN--WDQECTPEMAVT 264
Db 1190 NYPSNTPQOQDETGTGKPVDTVTGLPYDPSTGEIIPATKLPPIPSVAGDEIILEVLNIT 1249
QY 265 N-----VDVQ-----PSTEMEGLOS-----INP-----NATG-----ERR 290
Db 1250 TDEVYGLPIDLETGLPRDPSGLPOLPNCTLVDPSPNKKPIPGSHSGFINSTSGEQSHED 1309
QY 291 VLVGVPCKTGT--MMOPD-----AETGEFLMARDTNY-----QNM 323
Db 1310 PSTGKPLDPTGILHPDEDSGLINPFTGDKLQSGSTGFWPGKRGQENGIMTPEQI 1369
QY 324 IESIDENGIVYVNE-----DAI-----LKELDVEYVVCPTFLGG----- 357
Db 1370 LEAL--NKLPTSNENVISPRPSDAVDRPTNTWMKTSQTYOYDGKKTIIIGSAAVYIH 1427

```

```

QY 358 -----RDWPSAAL--NPDGSIYFIPLNNVCYDMAVDO-----EFTSMDYVNTSNYTK 403
Db 1428 TALGTPQTDPPTTGLPSPDSPTGIPPIPGFNVLADPQTEQIKGSPVYSLVYKEKNITTE 1487
QY 404 LPFGKDMIGRIDALIDISTGRTLSVERAANYSPLVSTG-----GVLFNCGTDRYFRA 457
Db 1488 AAYGL-----PVDPKTGPEPIDPI-----SYLPFAKNELIDPIISGKVFSGSIAGFLSG 1535
QY 458 -----LSOETGETLMQT-----RLATVASGQALSYEVGMQYVAIAG----- 494
Db 1536 KAGSOSKSSDESGNPIDPSTNMPYDPKGGKLLDPESGIALDINSVGV--FATVPGTAPARK 1594
QY 495 GGV 497
Db 1595 GGV 1597

```

Search completed: May 24, 2002, 10:20:13
 Job time: 326 sec

Fri May 24 11:27:29 2002

walick-934-128.pep.ra1

339 AILKELDVEYDCPTFLGGRDWPSSAALNPPDSGIYFLPLNNVC

Db 412 GYTLTGKFWYIGIPGLAHNFMGMAVSPRTHLVLPALQIPFGKNOVGSEKPHDPAN 471
 QY 386 VDEFTSMQVNTSNVTKLPKPKDMIGRIDIDISTGRTLSVEAANANSPILSTGGV 445
 Db 472 VGLDMTKNGLPDPE-ARVAYIKDLHGLWLLAMPVKMETVWKIDHKPMNGVLANGDL 530
 QY 446 LFNQGDYFRALSOETGELMOTRLATVAGSQAISYEVDGMOYA-----IAGG 495
 Db 531 LFGGLANGFEHAYDATNGSDLYKFDAGSGITIASPRTYSVNGKQYVAVEVGSGIYPSNG 590
 QY 496 GVSYSGS 502
 Db 591 GVGRTSG 597

RESULT 2

quinohemoprotein ethanol dehydrogenase (EC 1.1.99.-) type 1 precursor - Comamonas testos
 C:Species: Comamonas testos
 C:Date: 08-May-1995 #sequence, revision 21-Jul-1995 #text_change 02-Jun-2000
 C:Accession: S62366; S62373; S65308; S52317
 R:Stoerovogel, J.; Kraayveld, D.E.; van Sluis, C.A.; Jongejan, J.A.; de Vries, S.; Duine, J.
 Eur. J. Biochem. 235, 690-698, 1996
 A:Title: Characterization of the gene encoding quinohemoprotein ethanol dehydrogenase
 A:Reference number: S62366; MUID:96184549
 A:Accession: S62366
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-708 <STOI>
 A:Cross-references: EMBL:X81880; NID:9663195; PIDN:CAA57464.1; PID:9663196
 A:Accession: S62373
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 32-54 <STO2>
 R:de Jong, G.A.H.; Geerlof, A.; Stoerovogel, J.; Jongejan, J.A.; de Vries, S.; Duine, J.
 Eur. J. Biochem. 230, 899-905, 1995
 A:Title: Quinohemoprotein ethanol dehydrogenase from Comamonas testos. Purification
 A:Reference number: S65908; MUID:95324580
 A:Accession: S65908
 A:Molecule type: protein
 A:Residues: 32-50, 'X', 52-54, 477-483, 'X', 485-490 <DEJ>
 A:Experimental source: ATCC 15667
 C:Genetics: qheh
 C:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)
 F:1-31/Domain: signal sequence #status predicted <SIG>
 F:32-708/Product: quinohemoprotein ethanol dehydrogenase type 1 #status experimental <MA>

Query Match 17.8%; Score 481; DB 2; Length 708;
 Best Local Similarity 25.4%; Pred. No. 4,7e-27;
 Matches 150; Conservative 84; Mismatches 239; Indels 118; Gaps 19;

QY 3 PSLMASGALALLAFAFAVPTVDE-----DELANPP-AGEWISYQONEN 50
 Db 12 PGRWVWLLAACLG--SAAAFQTPPAQAAAANVORVGDGDIRANARPTDPTIGVDAE 69
 QY 51 YHSPDTQITTEVNGOLVYVARGMPCK-VQVTPLIHGVYMLANPGVIOAIDAKTGD 109
 Db 70 TRKSLDQINANAKDLGLAMSYNLESTRGVEATPVVGVIMYVASSSVYHAIDRTGN 129
 QY 110 LIWHRRLPP-----IA----- 122
 Db 130 RIMTYDIPDINSTGFKCCDVNRGVALMKKGVYVYAGDRLIALDAATGEVWHQTFE 189
 QY 123 -----TINSFGIYANGVIAGST-CQYSPFCFVSGHDSATGELMRNFEP----- 168
 Db 190 GOKSLITITGAPRYEKGKVIIRGAECVGRG-YITVADAFETGKRKRWISVGGDSKPE 248
 QY 169 --RAGEGDETWGNDYEAAMTGA-----WGQITVDPTNLVHYGSTAVGPASETORTPT 221

Db 249 EDESMKRAATW--DPSGKMEAGCGTMDMSMFEALNTMYVGNNGSPSHKRSR 306
 QY 222 GGLYGTNFRFAVRPDPTGEIVRHOPLPRNDDECFEEMAVTNVDVOSTMEGLQIN 281
 Db 307 GGNLVIASIVALLDPTGKYKMHQETGPDQMTTSTQPIILADIKI----- 353
 QY 282 PNATGERVVLGVCCKTGWQFDATGEFLNARDTNYQNIESTIDENG-LVYNENAI 340
 Db 354 ---ACKPRVITLHAD-KNGFEFLDRTNGKFIKAFNEFVPMNASGDKHGRIGI---AA 406
 QY 341 LKELVEYDVCPTFLGGRDWPSSALNPDGIVPIPLNNCYMAVDO-EFT----- 391
 Db 407 ARDGSKPDQDAVGPYPYGAHNMHPSFNPQGLVYLPANVPVLMDDKMEFNQAGPKPQ 466
 QY 392 SMQVNTSNVTKL-PPKDMIGRIDIDISTRTLSVERAANANSPILSTGGVLENG 450
 Db 467 SGTGMNTAKFEPNPEKSPFGRLAMPVQAOKAMSEHVSPPNNGTILTTAGNVYFOGT 526
 QY 451 TDRFRALSOETGELMOTRLATVAGSQAISYEVDGMOYAIA-GGVSYSG 500
 Db 527 ADGRLVYHAATEKLEAPGTGVVAPSTYIMDGROYVSAVAGWGVYG 577

RESULT 3

alcohol dehydrogenase, membrane-bound (EC 1.1.-.-) 72k chain precursor - Acetobacter
 C:Species: Acetobacter polyoxogenes
 A:Variety: strain NB11028
 C:Date: 21-Nov-1993 #sequence, revision 26-Jul-1996 #text_change 20-Jun-2000
 R:Tamakl, T.; Fukaya, M.; Takemura, H.; Tayama, K.; Okumura, H.; Kawamura, Y.; Nishiy
 Biochim. Biophys. Acta 1088, 292-300, 1991
 A:Title: Cloning and sequencing of the gene cluster encoding two subunits of membrane
 A:Reference number: S14270; MUID:91159482
 A:Accession: S14270
 A:Molecule type: DNA
 A:Residues: 1-738 <TAM>
 A:Cross-references: GB:D00635; NID:9216185; PIDN:BAA00528.1; PID:9216186
 A:Experimental source: strain NB11028
 C:Complex: heterodimer of 72k and 44k chains
 C:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)
 C:Keywords: alcohol metabolism; membrane protein; oxidoreductase
 F:1-35/Domain: signal sequence #status predicted <SIG>
 F:36-738/Product: alcohol dehydrogenase 72k chain #status predicted <MA>

Query Match 17.7%; Score 479; DB 2; Length 738;
 Best Local Similarity 25.9%; Pred. No. 7e-27;
 Matches 153; Conservative 77; Mismatches 230; Indels 130; Gaps 16;

QY 10 SAGALALLAFAFAVPTVDE-----DELANPP-AGEWISYQONENYRSPITOTT 61
 Db 17 TAGTICALLISGYATMASADGCGATGEALIHADHPGMMYGRYSDGRSPLDQINR 76
 QY 62 ENVGOLQVWARGMPCKVQV-TPLIHGVYMLANPGVIOAIDAKTGDILWHRRLPP- 119
 Db 77 SNVGNLKLAWLIDTNRGEGTPIVDGMVYATTNWMSKMAVDAATCKLMSIDPRAVG 136
 QY 120 NIA-----TINS-----FGI----- 129
 Db 137 NIADGCGDVTNNGAAVWNGKYVFGFDSRLALDAKTGKLWMSVNTIPPEAELGRQSY 196
 QY 130 -----VANGVIVAGSTQYSPFC--FVSGHDSATGEELMRNYETIPRAGEGD----- 175
 Db 197 TVDGAIRIAKRIIION--GSEFGARGFVSFAFETGKYDMREFYVTPNKKEPDASDS 254
 QY 176 -----ETWGDYEAAMTGA-----WGQITVDPTNLVHYGSTAVGPASETOR 218
 Db 255 VLMNKAAYQTS-----PTGAMTRQGGGQVWDSIYDPAVLVLYGNGSPNNYKXR 307
 QY 219 GTPGGLYGTNFRFAVRPDPTGEIVRHOPLPRNDDECFEEMAVTNVDVOSTMEGLQ 278
 Db 308 SEGKGNLFLGSIYALKEPGEYVWHQETGPDQMTTSTQPIILADIKI----- 357

QY 281 NENATGE-RRVLTGVCPTKTMQFQDAEGERFARPTNQNMTESID-----EN 330
 Db 339 -----GEMRHVYHAP-KNGEFTYVADAKTGEISGKRYVYONNANGIDPLTRGPNPD 411
 QY 331 GIVTVNDDALLKELDEYDYCPFTLGGKDMPSALNPDSCGYFTPLNNVCY----- 381
 Db 412 GLYTLNG-----KFWYGI-PGPIGAHNFMAMAYSPKTHLYIPAIQIFGCKNOVG 463
 QY 382 -----DMWAVDQFTSMQVNTSNVTKLPDPRGDKMGRDAIDISTGRTLSVERAANYS 437
 Db 464 KPRADSNWVGLDWTKNGLPDPPE-AKTAVIDLHGLMLADPVKMETWVKIDHKGPWNG 522
 QY 438 VLTGGVLFNGGTDRTYPRALSQETGFTLMQTRLATVASQAISEYEDGMOYVA----- 491
 Db 523 ILATGDLPLFGLANGEEHAYDATNGSDLYKFDQSGIAPMYFTSVNKKOYVAEVGWG 582
 QY 492 -----IAGGVSYSG 502
 Db 583 GIVPISMGVGTGSG 597

RESULT 6

J00706
 N:Alcohol dehydrogenase (acceptor) (EC 1.1.99.8) alpha chain precursor - Methylobacterium
 C:Species: Methylobacterium extorquens
 C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 08-Oct-1999
 R:Accession: J00706; S07908
 R:Anderson, D.J.; Morris, C.J.; Nunn, D.N.; Anthony, C.; Lidstrom, M.E.
 A:Title: Nucleotide sequence of the Methylobacterium extorquens AM1 *mox* and *mox* genes
 A:Reference number: J00706; MUID:90337342
 A:Accession: J00706
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-626 <AND>
 A:Cross-references: GB:M31108; NID:q150017; PIDN:AAA25380.1; PID:q150018
 R:Experimental source: strain AM1
 R:Nunn, D.N.; Day, D.; Anthony, C.
 Biochem. J. 260, 857-862, 1989
 A:Title: The second subunit of methanol dehydrogenase of Methylobacterium extorquens AM1
 A:Reference number: S04644; MUID:89350892
 A:Accession: S07908
 A:Molecule type: protein
 A:Residues: 28-50, 'XX', 53 <NUN>
 A>Note: The source is designated as Methylobacterium extorquens AM1
 C:Comment: This enzyme oxidizes methanol to formaldehyde.
 C:Genetics:
 A:Gene: *mox*
 C:Keywords: alcohol metabolism; oxidoreductase
 F:1-27/Domain: signal sequence #status predicted <SIG>
 F:28-626/Product: alcohol dehydrogenase (acceptor) #status experimental <MAT>

Query Match 16.0% Score 434; DB 2; Length 626;
 Best local similarity 24.5%; Pred. No. 1e-23;
 Matches 148; Conservative 91; Mismatches 208; Indels 158; Gaps 23;

QY 10 SAAALMLA-AP-AFAQVTPYDELAPNPAE-WISYQNOENYRHSPLTQITENNQ 66
 Db 7 SVALAMLAALAPALSCGAYANDKVELSKDDWVWPKNYSNPFSLDKOIKNGVNO 66
 QY 67 LQLVW-ARGMQPKVYVTPILHGVW-----LANPQV----- 99
 Db 67 LRPAMTFSTGLNGH-EGAPLVVDGKMYIHTSPNNTFALGIDDPETILMOKPKQNPAA 125
 QY 100 -----IOAIKAGTGLIWEHRLPNIATLN 125
 Db 126 RAVVACCDLVNGLAYWPGDGKTPALIKTQDGNVAALAAETGVTM----- 173
 QY 126 SFGIVANGVIAGSTGQSPF-----GC-----FVSGHDSATGELMRNFEI- 167
 Db 174 -----VENSIDIKVSTLTIAPIYVKRVIIIGSSGAEIGVAGYLAVDVKTSGQVRAVATG 229

QY 168 -----PRAGEE--GDETWGNDYERAMTGA---WGOITYDPVTLNHYGS 207
 Db 230 PCKDLLASDENIKNPHGQGLGTGWEGD---AMKIGGCTMMGVAVADPSTNITLFT 286
 QY 208 TAVPASETQRTPGTLYGNTREFAVRPPDGEIYWHOTLPDRNMDQCTEMAVTV 267
 Db 287 GNPAPNNEETMR--PGDNKK-TMIFGRDADTGAKRGYQATPDEWD-----YACVN 335
 QY 268 VQSTEMEGLDSINPNATGERRVLTVGCTGTMMQFPAETGEFLMAR---DNYOMI 324
 Db 336 VAMLEQND-----KDSARKLLTHPDRNGIYVTLTRDGLALVSNKDDI--VAVF 385
 QY 325 ESIDENGIVTNEEDALLKELD-VEYDVCPTFLGGRMPSALNPDSCGYFTPLNNVCY 383
 Db 386 KSVLDKTCQPVPRDEYGTDRMDHLAKDICSAMGYNQCHSDYDPKRELFEGINHLICMD 445
 QY 384 MAVDQFTSMQVNTSNVTKLPDPRG-----DMIGRIDAIDISTGRTLSVERAANYS 437
 Db 446 EPMFLPYRAGQFVGAATLNNYPPGPKGRQNEGLQIKAVNATGDYKMEKMRFAVWG 505
 QY 438 VLTGGVLFNGGTDRTYPRALSQETGFTLMQTRLATVASQAISEYEDGMOYVAIGGV 497
 Db 506 TMAIAGDLVEYGTLDGYLAKRDSQDGLLMFKIPSGAIGTPMYTHKGTQYVAI----- 560
 QY 498 SYSGS 502
 Db 561 YGVG 565

RESULT 7

B83399
 quinioprotein alcohol dehydrogenase PA1982 [imported] - Pseudomonas aeruginosa (strain
 C:Species: Pseudomonas aeruginosa
 C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
 R:Accession: B83399
 R:Stover, C.K.; Pham, X.O.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.;
 adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lardig, K.; L
 Nature 406, 959-964, 2000
 A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pa
 A:Reference number: A82950; MUID:20437337
 A:Accession: B83399
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-623 <STO>
 A:Cross-references: GB:AE004624; GB:AE004091; NID:9947973; PIDN:AA05370.1; GSPDB:GN
 A:Experimental source: strain PA01
 C:Genetics:
 A:Gene: *exaA*; PA1982

Query Match 15.4% Score 417.5; DB 2; Length 623;
 Best local similarity 23.2%; Pred. No. 1.7e-22;
 Matches 139; Conservative 94; Mismatches 248; Indels 119; Gaps 19;

QY 3 PPSLIMASAGALAL---IAAPFAQVTPYDELAPN-PPAEWISYQNOENYRHSPLT 57
 Db 9 PAGLRPSLHCLAFVALGSAALAKDVTMEDIAKDQTTDVTQYQMGTHAQRMSPLK 68
 QY 58 QTTTNGCOLLVVARGM-----QPGKQVTPILHGVWVLANPQVIOAIDAKTGLIME 113
 Db 69 QVNDADNVEKTLPMMSYSFGEKGRQ-ESQAIYSDGVIYVAYSSRLPALDAKTKGKLT 127
 QY 114 HRQLPN-----LATLNSGCIAN---GVIV-----AGST 140
 Db 128 YNHLPPDIPRCCDVNNGAATYGKVFEGTIDASVYALNNKGVKPKFADHAGYT 187
 QY 141 QVYSP-----FGCF--VSGHDSATGELMRNFEI----- 169
 Db 188 MTGAPITVADGKTKVLLHSSGDEFGVGRFLARPDPDGEELIMRPFEGHMGKLNK 247
 QY 170 -----AGEBDEFTWMDYEA-----RMWTGA---WGOITYDPVTLNHYGSTAVGASSET 216

QY 314 WARDTN-YONMIESIDENGIVIVNEEDRI LNEEDVZ

A;pathway: respiratory chain

1

C:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)
 C:Keywords: oxidoreductase; pyrroloquinoline quinone; respiratory chain; transmembrane F
 F:11-37/Domain: transmembrane #status predicted <TM1>
 F:41-59/Domain: transmembrane #status predicted <TM2>
 F:63-81/Domain: transmembrane #status predicted <TM3>
 F:96-110/Domain: transmembrane #status predicted <TM4>
 F:120-140/Domain: transmembrane #status predicted <TM5>
 F:93-95/Binding site: transmembrane #status predicted <TM5>
 F:466/Active site: ubiquinone (Arg, Asp) #status predicted

Query Match 12.6%; Score 342; DB 1; Length 796;
 Best Local Similarity 22.7%; Pred. No. 7,5e-17;
 Matches 160; Conservative 74; Mismatches 202; Indels 270; Gaps 32;

```

6 LMAS-----AGALADLAAPAFQVPTVDELNANPAGEMWISGONENYRHSPLTQ 58
135 LTMAGFNDPQELNGTISADATPAEA-ISPVADQ-----DMPAYGRNDEGGRFSPKQ 185
59 ITTENVGOLQLVMA-----RGMOPKV--QVPLIHGVMYLANPGDVIAIDAKTGD 109
186 INADVHNLEKAWFRTGDKVQPNPGEITNEVTPIKVGYDTLYCTAHORLEALDASGK 245
110 LIWEHRROLPIATLNSF-----GI-----VANGYVA- 137
246 EKMHD--PELKTNESFOHTCGVSYHEAKAETASPEVADCPRIILPVNDGRILAI 302
138 ---GSTCQ-----YSP-----FG 147
303 NAENGKLCETFAKGVNLQSNMPTKPGLYEPTSPITITDKTIYAGSVTDNFSRETSS 362
148 CFVSGHDSATGEELMKNYFIPRA-----GEGDETCNDYEARMGTGANGQITTYDPT 200
363 GYIRGFVNTGELLMA--FDGAKDPNALPSDEHTFTNS-----PNSWAPAAYDAKL 413
201 NLVHGSTAVGPASETQRTGPGGLYGTNRFAVRPDTSEIYVHRQTLPRNMDECTPE 260
414 DIVLPLMGVTPPDINGNKRTPQDERYASSI-LALNATGTLKMSYQVYHDLMDM-- 468
261 MMTVNDVOPSTMEGLQSLNPNATGERRYLVGPKCTGTMMOFDETEFTL----- 313
469 ---LPAQPT---LADITVN--GQKVPYIAPAKTGNIFVLDNRNGLVPAPEKRP 515
314 ---WARTNYQNNIESIDENGI-----VTNEDAILKEL--DVEYD----- 349
516 VPGAAKGDVYTPQPSSELSFRPTKLSGADMGATMFDQLVCVRFHQRKEGIFTPP 575
350 ---VCPFLGGRDWPSSALNP----- 367
576 SEQGLVFPGNLGMFENGISVDPNREVAIANPMALPFVSKLIPRGPNRMEQPKDAKGT 635
368 ---DSGT-----YFIPLNNVCYDMAVADQFTSMQVNTSNVTKLPKGMIGRIDAID 418
636 GTEGSIQPOGYGVYVTLN-----PFLS-----PGLPCQKQPMAGYISALD 676
419 ISTGRTLSVERAANYS-----PV-----ISTGGVLFENGCT--DRFERAL 458
677 LKTNEVVMKKRIGTPQDSMFPMPVPFPMNGMPLGGLISTAGNVLFIAATADNLRAY 736
459 SQETGETLMQTRLATVASQA--ISTEVDGMQYVAI--AGGGSYGS 501
737 NMSGKELMQRGLP--AGGQATPMTEYVNGKQYVVISAGHGSGFT 780

```

RESULT 10
 H90644
 glucose dehydrogenase [imported] - Escherichia coli (strain O157:H7, substrain RIMD 0509
 C:Species: Escherichia coli
 C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Aug-2001
 C:Accession: H90644
 R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
 gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
 DNA Res. 8, 11-22, 2001

A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and 9
 A:Reference number: A95629; MUID:21156231; PMID:11258796
 A:Accession: H90644
 A:Status: Preliminary
 A:Molecule type: DNA
 A:Residues: 1-796 <HAY>
 A:Cross-references: GB:BA000007; PIDN:BBJ33551.1; PID:q13359584; GSPDB:GN00154
 A:Experimental source: strain O157:H7, substrain RIMD 0509952
 C:Genetics:
 A:Gene: EC50128
 C:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)

Query Match 12.6%; Score 342; DB 2; Length 796;
 Best Local Similarity 22.7%; Pred. No. 7,5e-17;
 Matches 160; Conservative 74; Mismatches 202; Indels 270; Gaps 32;

```

6 LMAS-----AGALADLAAPAFQVPTVDELNANPAGEMWISGONENYRHSPLTQ 58
135 LTMAGFNDPQELNGTISADATPAEA-ISPVADQ-----DMPAYGRNDEGGRFSPKQ 185
59 ITTENVGOLQLVMA-----RGMOPKV--QVPLIHGVMYLANPGDVIAIDAKTGD 109
186 INADVHNLEKAWFRTGDKVQPNPGEITNEVTPIKVGYDTLYCTAHORLEALDASGK 245
110 LIWEHRROLPIATLNSF-----GI-----VANGYVA- 137
246 EKMHD--PELKTNESFOHTCGVSYHEAKAETASPEVADCPRIILPVNDGRILAI 302
138 ---GSTCQ-----YSP-----FG 147
303 NAENGKLCETFAKGVNLQSNMPTKPGLYEPTSPITITDKTIYAGSVTDNFSRETSS 362
148 CFVSGHDSATGEELMKNYFIPRA-----GEGDETCNDYEARMGTGANGQITTYDPT 200
363 GYIRGFVNTGELLMA--FDGAKDPNALPSDEHTFTNS-----PNSWAPAAYDAKL 413
201 NLVHGSTAVGPASETQRTGPGGLYGTNRFAVRPDTSEIYVHRQTLPRNMDECTPE 260
414 DIVLPLMGVTPPDINGNKRTPQDERYASSI-LALNATGTLKMSYQVYHDLMDM-- 468
261 MMTVNDVOPSTMEGLQSLNPNATGERRYLVGPKCTGTMMOFDETEFTL----- 313
469 ---LPAQPT---LADITVN--GQKVPYIAPAKTGNIFVLDNRNGLVPAPEKRP 515
314 ---WARTNYQNNIESIDENGI-----VTNEDAILKEL--DVEYD----- 349
516 VPGAAKGDVYTPQPSSELSFRPTKLSGADMGATMFDQLVCVRFHQRKEGIFTPP 575
350 ---VCPFLGGRDWPSSALNP----- 367
576 SEQGLVFPGNLGMFENGISVDPNREVAIANPMALPFVSKLIPRGPNRMEQPKDAKGT 635
368 ---DSGT-----YFIPLNNVCYDMAVADQFTSMQVNTSNVTKLPKGMIGRIDAID 418
636 GTEGSIQPOGYGVYVTLN-----PFLS-----PGLPCQKQPMAGYISALD 676
419 ISTGRTLSVERAANYS-----PV-----ISTGGVLFENGCT--DRFERAL 458
677 LKTNEVVMKKRIGTPQDSMFPMPVPFPMNGMPLGGLISTAGNVLFIAATADNLRAY 736
459 SQETGETLMQTRLATVASQA--ISTEVDGMQYVAI--AGGGSYGS 501
737 NMSGKELMQRGLP--AGGQATPMTEYVNGKQYVVISAGHGSGFT 780

```

RESULT 11
 H85495
 glucose dehydrogenase [imported] - Escherichia coli (strain O157:H7, substrain EDL933
 C:Species: Escherichia coli
 C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
 C:Accession: H85495
 R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; May

iller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamocousis, K.; Apodaca, Nature 409, 529-533, 2001
 A:Title: Genome sequence of enterohemorrhagic *Escherichia coli* O157:H7.
 A:Reference number: AB5480; MUID:21074935; PMID:11206531
 A:Accession: H85495
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-796 <STO>
 A:Cross-references: GB:AE005174; NID:q12512839; PIDN:AG54428.1; GSPDB:GN00145; UWGP:201
 A:Experimental source: strain O157:H7, substrain EDL533
 C:Genetics
 A:Gene: gcd
 C:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)

Query Match 12.6%; Score 342; DB 2; Length 796;
 Best Local Similarity 22.7%; Pred. No. 7.5e-17;
 Matches 160; Conservative 74; Mismatches 202; Indels 270; Gaps 32;

QY 6 LIMS-----AGALALAPAFVAVTVDLAPNAPAGEMTSYGONENRHSPLTQ 58
 DB 135 LTMAGFNDPQELNGLTADATPAEA-ISPVAOQ-----DMPVGRNDSGCKSPFLKQ 185
 QY 59 ITTENVGOLQVMA-----RGMOPGV--GVTPLIHDGVYLANPQDVIOAIDATGCD 109
 DB 186 ITADVHNLKEAWFRTGDKOPNDPEITNEVTPIKVGDLYLCTAHQRFLALDAASGK 245
 QY 110 LIMEHRQLPNIATLNSF-----GI-----VANGVIA- 137
 DB 246 EKMHPD-----PELKTWESFOHTVCRGVSYHEAKETASPEVADCPRIILPNDRLIAI 302
 QY 138 ----GSTCQ-----YSP-----FG 147
 DB 303 NAENGKICETFAFKVGLNLSQNMPTKRGLEPTSPILIDKTIYAGSVTDNSTRFS 362
 QY 148 CFVSGHDSATGELMKNYFIPRA-----GEDEDMGNDYFAWMGTGACQIITYDPT 200
 DB 363 GVIRGFDVNSGKLLMA--FDPGAKDPNAPISDEHTFTNS-----PNSWAPAAADAKL 413
 QY 201 NLVHGSTAVGPASSTORCGTGGTLYGTNTREAVRBDTGEIYWRHQLLPNDNDDECTFE 260
 DB 414 DLVYLPNGVTPPDIMGNKRTPEQERYASSI--LALNATTKLAMSXYVHDLMDMD----- 468
 QY 261 MNVTVNDVQSTEMEGLOSINPNAATGERRYLVGPCKTGTMOPFAETGEF----- 313
 DB 469 ----LPAQPT-----LADITV--GQKVPYIYAPAKTGNIFVLDRNGELVYPAPEKP 515
 QY 314 ----WARDNINQNMIESIDENGI-----VTNEDAILIKEL--DVEYD----- 349
 DB 516 VPQGAAGGDTYTPQPESELSFRPTKDLGADMMGATMFDOLVCRVMEHOMTBEITFP 575
 QY 350 ----VCPTLGRDMPSSALNP-----D----- 367
 DB 576 SEGGTLVFPNGINLMEFEGISVDPREVAIANPMLAPFVSKLIPRGPNMPOPKDAKT 635
 QY 368 --DSGI-----YFPLNNVCYDMMAVDQETSMVYNTKLPCKDMIGRIDAID 418
 DB 636 GTSBGLOPQYGVYGTLLN-----PFLS-----PGLPCQKQPAAGYSALD 676
 QY 419 ISTGRILSVYERAAATIS-----PV-----ISTGGVLENGGT--DRYFRL 458
 DB 677 LKTNEVYVWKKRIGTPODSMPFPVYVFPNMGMPMLGSPISTAGNVLFIAIADNYLRY 736
 QY 459 SQEGETLMQTRLATVASSQA--ISTEYDGMQYVAI--AGGVSYS 501
 DB 737 NMSGKILMGRLP--AGGQATPMTEYVNGKQYVVISAGHGSFGT 780

RESULT 12
 AG0523
 glucose dehydrogenase [imported] - Salmonella enterica subsp. enterica serovar Typhi (st
 C:Species: Salmonella enterica subsp. enterica serovar Typhi
 A:Note: this species has also been called Salmonella typhi

C:Date: 09-Nov-2001 #sequence, revision 09-Nov-2001 #text, change 27-Nov-2001
 C:Accession: AG0523
 R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Church
 T.; Khoule, S.; O'Garra, P.
 Nature 413, 848-852, 2001
 A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens,
 A:Title: Complete genome sequence of a multiple drug resistant *Salmonella enterica* se
 A:Reference number: AB0502; PMID:11677608
 A:Accession: AG0523
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-796 <PAR>
 A:Cross-references: GB:AL513382; PIDN:CAD01327.1; PID:q16501455; GSPDB:GN00176
 C:Genetics
 A:Gene: STY0191
 C:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)

Query Match 12.1%; Score 327; DB 2; Length 796;
 Best Local Similarity 22.7%; Pred. No. 9.3e-16;
 Matches 157; Conservative 68; Mismatches 226; Indels 240; Gaps 31;

QY 6 LIMS-----AGALALAPAFVAVTVDLAPNAPAGEMTSYGONENRHSPLTQ 58
 DB 135 LTMAGFNDPQELNGLTADATPAEA-ATSSIADE-----DMPVGRNDSGCKSPFLKQ 185
 QY 59 ITTENVGOLQVMA-----RGMOPGV--GVTPLIHDGVYLANPQDVIOAIDATGCD 109
 DB 186 ITADVHNLKEAWFRTGDKOPNDPEITNEVTPIKVGDLYLCTAHQRFLALDAASGK 245
 QY 110 LIMEHRQLPNIATLNSF--GIANGV-----IVA----- 137
 DB 246 EKMHPD-----POLKTSFQHTVCRGVSYHEAKADTASPEVADCPRIILPNDRLIAI 302
 QY 138 ----GSTCQ-----YSP-----FG 147
 DB 303 NAENGKICETFAFKVGLNLSQNMPTKRGLEPTSPILIDKTIYAGSVTDNSTRFS 362
 QY 148 CFVSGHDSATGELMKNYFIPRA-----GEDEDMGNDYFAWMGTGACQIITYDPT 200
 DB 363 GVIRGFDVNSGKLLMA--FDP--GAKDPNTIPABEHAFTNSPNSMAPAAYDAKDLVYL 418
 QY 206 GSTAVGPASSTORCGTGGTLYGTNTREAVRBDTGEIYWRHQLLPNDNDDECTFE 265
 DB 419 PMGVTTPDIMGKRTPEQERYASSI--LALNATTKLAMSXYVHDLMDMDLPDRAOPTLAD 477
 QY 266 VDVOPSTEMEGLOSINPNAATGERRYLVGPCKTGTMOPFAETGEF----- 312
 DB 478 ITVDGTV-----PVITAPAKTGNIFVLDRNGELVYPAPEKP 520
 QY 313 ----WARDNINQNMIESIDENGI-----VTNEDAILIKEL--DVEYD----- 349
 DB 521 AKGDYVAKTOPFSDLTRPRKKDLGADMMGA--TMDOLVCRVMEHOMTBEITFP 579
 QY 340 ILKELDEYDVCPFTLGRDMPSSALNPDSGIY-----IPLNNVCY-----D 382
 DB 580 TL-----VFPGLGFEWGISVDPQVAVIANPMLAPFVSKLIPRGPNMPOPKDAKT 631
 QY 383 MNVTVNDVQSTEMEGLOSINPNAATGERRYLVGPCKTGTMOPFAETGEF----- 312
 DB 632 AKGTGTAGIOPQYGVYGTLLN-----PFLS-----PGLPCQKQPAAGYSALD 676
 QY 434 NYS-----PV-----ISTGGVLENGGT--DRYFRL 458
 DB 692 RDSMPFPVYVFPNMGMPMLGSPISTAGNVLFIAIADNYLRY 736
 QY 474 VASGQA--ISTEYDGMQYVAI--AGGVSYS 501
 DB 751 --AGGQATPMTEYVNGKQYVVISAGHGSFGT 780

RESULT 13

500943
 glucose dehydrogenase (pyrroloquinoline-quinone) (EC 1.1.99.17) - Acinetobacter calcoace
 C:Species: Acinetobacter calcoacetius
 C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
 C:Accession: S00943
 R:Cleaton-Jansen, A.M.; Goosen, N.; Odle, G.; van de Putte, P.
 Nucleic Acids Res. 16, 6228, 1988
 A:Title: Nucleotide sequence of the gene coding for quinoprotein glucose dehydrogenase
 A:Reference number: S00943; MUID:88289368
 A:Accession: S00943
 A:Molecule type: DNA
 A:Residues: 1-801 <CDS>
 A:Cross-references: EMBL:X07235; NID:938711; PIDN:CAA30222.1; PID:g38712
 A:Experimental source: strain LMD 79, 41
 C:Function:
 A:Description: catalyzes the oxidation of D-glucose to D-gluconic acid by ubiquinone
 A:Pathway: respiratory chain
 C:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)
 C:Keywords: oxidoreductase; pyrroloquinoline-quinone; respiratory chain; transmembrane
 F:9-35/Domain: transmembrane #status predicted <TM1>
 F:39-57/Domain: transmembrane #status predicted <TM2>
 F:61-79/Domain: transmembrane #status predicted <TM3>
 F:94-108/Domain: transmembrane #status predicted <TM4>
 F:118-137/Domain: transmembrane #status predicted <TM5>
 F:91.93/Binding site: ubiquinone (Arg, Asp) #status predicted
 F:471/Active site: Asp #status predicted

Query Match 12.0%; Score 323.5; DB 1; Length 801;
 Best Local Similarity 21.8%; Pred. No. 1.7e-15;
 Matches 152; Conservative 75; Mismatches 193; Indels 277; Gaps 31;

20 PAFAQVPTVDLNPARGENISTGQONERYRSPLOTITTEVGGLOLVNARGQPK 79
 152 PETAQAVPGVA-----SDMPAGRTQAGVRSPLKQINDQVNDLKAAT--LRGTD 202
 80 V-----QVPLIHGVMYLANPGDVIOADAKTDLWEHRRLPNIATLNSF- 127
 203 LKTDNDSEETNQVTPRIKGNMFCIAHQOLAIIDPATGKEM--RDPPLKTKDSFQ 259
 128 ----- 127
 260 HLTGQVWYDANNTEPAISLSQKSSSTQCPKVFVENDGRILVANAADGKACDFG 319
 128 -----GIYANG-VIVAGS--TCQSS--PGCVSCHSDATG 158
 320 ONGOVNIOEFPYAPGPGYPTSGIYVTSYVAGSVTDNISKPSG-VIRGYDVNTG 378
 159 EELRNNTFIPRAGE-----EGDETFGNITYEARMTGAMGOITDVPVNLVHYGSTAVP 212
 379 KILW--VEDTGADPNAMRGEGETFEVHNS-----PNMAAPLAYDAKLIV-VYPTGV-- 427
 213 ASETQRTPGGLYG-----TNTREAVRPDTGIVRHOITLPRDWDDECTFEEM 262
 428 -----GTP--DIWGDRTLEKERYANSMIAINSTGLVANNPQTHHLDMDVPSOPS 479
 263 VTNVDVOPSTMEGLQINP--NAATGERVLTGV-----CTTG--TWMOFD 306
 480 LADIKRKAQGVPAIYVLTGKNAFVLDNR--NGQPTVPTKRPQVTKRGQKGEY 537
 307 AETGER-----LMA-----ROTNQNMLESIDENGIYTVNED 338
 538 SKTQPSDINLAPODKLIDKDMGATMLDQLMCRVSFKRLNTDGIYTPRENGTL----- 592
 339 AILKEIDVEYDPCPTFLGGRDMPASALNPS-----GIYFI----- 374
 593 -----VFPGNLGVFEWGMGSVNDPROVAVNPIGILPEFYSRLIPADPNRAQAKG 641
 375 -----PLANNCYDMAVADEFTSMADVNTSNVTKLPPCGMDMIGRIDAIDISTGRTLW 426
 642 AGTEGQVQPMGVY-----GVEISATLSPLGLPCQOPAMGVYAGADLKTREYVW 691
 427 -----SVERAANYSPVLSGGGVLENGGT--DRYFRALISQETGLLW 467

Db 692 KKRIGTIRSLPRLPQPAVKIGVPGCGSISTAGNWFVGATQDNLVRAFNVTNKRKL 751
 QY 468 QTRLAIVASQA--ISYEVDGMQVAVI-AGGVSYS 501
 Db 752 EARLP--AGGQATPMTYEINGQYVIMAGGHSFGT 786

RESULT 14

QPKEX

glucose dehydrogenase (pyrroloquinoline-quinone) (EC 1.1.99.17) - Gluconobacter oxyda
 C:Species: Gluconobacter oxydans
 C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 22-May-1998
 C:Accession: S17716; S19265
 R:Cleaton-Jansen, A.M.; Dekker, S.; van de Putte, P.; Goosen, N.
 Mol. Gen. Genet. 229, 206-212, 1991
 A:Title: A single amino acid substitution changes the substrate specificity of quinop
 A:Reference number: S17716; MUID:92017653
 A:Accession: S17716
 A:Molecule type: DNA
 A:Residues: 1-808 <CDS>
 A:Cross-references: EMBL:X62710
 R:Goosen, N.
 submitted to the EMBL Data Library, February 1992
 A:Reference number: S19265
 A:Accession: S19265
 A:Molecule type: DNA
 A:Residues: 1-212, 'A', 214-808 <GOO>
 A:Cross-references: EMBL:X62710; NID:958416; PID:958417
 C:Genetics:
 A:Gene: gdh
 C:Function:
 A:Description: catalyzes the oxidation of D-glucose to D-gluconic acid by ubiquinone
 A:Pathway: respiratory chain
 C:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)
 C:Keywords: oxidoreductase; pyrroloquinoline-quinone; respiratory chain; transmembran
 F:9-28/Domain: transmembrane #status predicted <TM2>
 F:35-54/Domain: transmembrane #status predicted <TM3>
 F:60-76/Domain: transmembrane #status predicted <TM4>
 F:94-110/Domain: transmembrane #status predicted <TM5>
 F:122-138/Domain: transmembrane #status predicted <TM6>
 F:91.93/Binding site: ubiquinone (Arg, Asp) #status predicted <TM5>
 F:470/Active site: Asp #status predicted

Query Match 10.9%; Score 295; DB 1; Length 808;
 Best Local Similarity 20.1%; Pred. No. 2e-13;
 Matches 141; Conservative 83; Mismatches 230; Indels 248; Gaps 25;

13 ALALLAAPFAQVPTVDLNA-----NPPAGEISYSGQONERYRSPLOTIT 61
 128 AVIALFASLFTDPRHDSIGSELPQIANASPADPDVPASEHAGKQACDRNSPINOINA 187
 62 ENVGLOLWV-----ARGMQPKV--QVTPLIHGVMYLANPGDVIOADAKTGDILW 112
 188 TVSNILKVAWHIHKDKMANSNDPOQNTNPTPEFNNTLYMCSLHQKLEAVDAGATGVNWK 247
 113 EHRRL--PNIATLNSFGI-----YANGVYV----- 136
 248 VYDFRLQINPQFHLTCRGVSFHEPTRANMDSQNPAPTDCAKOSILEVNDGRILEVDAD 307
 137 AGSTC-----QYSPGCV----- 150
 308 TGTCTSGFNGNGETDLRVNQPPTTGGQLEPSPVITDKLIANSATDNGSVKQASGA 367
 151 -SGHDSATGELW-----RNYFTPRAGEGDETWGNDYEARMTGAMQOITDVPVNLVHY 205
 368 TQAFDYVTGKRVVFDASNDPNOPLPDESHPVFPNPNKSHYS-----SYDANLNLVYI 422
 206 GSTAVGPASSETQGRP-----GGTLVGTNRF-----AVRPDTGIVRHOITLPRDMDQDC 257
 423 PMGV-----GTPDQMGDRKTDSERFAPGIVALNADTGLAMFVGTVAHHDMDHEL 473

Fri May 24 11:27:30 2002

wallick-934-128.pep.rpr

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2002, 10:30:52 ; Search time 28.06 Seconds
(without alignments)
696.842 Million cell updates/sec

WALICK-934-128.PEP

Title: 2705
Sequence: 1 MKRISLMAAGALALAAP.....GMOYVIAAGGVSYGSLINS 505

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SWISSPROT_40.*

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	483.5	17.9	739	1	DHET_ACEBU
2	479	17.7	738	1	DHET_ACEBU
3	477.5	17.7	742	1	DHET_ACEBU
4	466	17.2	757	1	DHET_GLUOX
5	435	16.1	600	1	XOXF_PARDE
6	434	16.0	626	1	DHET_METEX
7	434	16.0	631	1	DHET_METEX
8	422.5	15.6	631	1	DHET_PARDE
9	417.5	15.4	631	1	DHET_PARDE
10	393	14.5	571	1	DHET_METEX
11	342	12.6	796	1	DHET_ACEBU
12	323.5	12.0	801	1	DHET_ACEBU
13	295	10.9	808	1	DHET_GLUOX
14	253	9.4	790	1	DHET_ACEBU
15	231.5	8.6	799	1	DHET_ACEBU
16	128	4.7	827	1	DHET_ACEBU
17	126	4.7	799	1	DHET_ACEBU
18	118.5	4.4	593	1	DHET_ACEBU
19	116	4.3	353	1	DHET_ACEBU
20	115.5	4.3	353	1	DHET_ACEBU
21	115.5	4.3	353	1	DHET_ACEBU
22	115	4.3	353	1	DHET_ACEBU
23	115	4.3	353	1	DHET_ACEBU
24	114.5	4.2	353	1	DHET_ACEBU
25	114	4.2	353	1	DHET_ACEBU
26	110.5	4.1	353	1	DHET_ACEBU
27	110.5	4.1	353	1	DHET_ACEBU
28	108.5	4.0	353	1	DHET_ACEBU
29	106.5	3.9	353	1	DHET_ACEBU
30	106	3.9	353	1	DHET_ACEBU
31	105	3.9	353	1	DHET_ACEBU
32	105.5	3.8	353	1	DHET_ACEBU
33	103.5	3.8	353	1	DHET_ACEBU

34	103.5	3.8	974	1	PHS2_SOLTU
35	103.5	3.8	1365	1	GFES_STRDO
36	102	3.8	595	1	YER5_SCHPO
37	102	3.8	790	1	PLMN_PIG
38	102	3.8	1323	1	MRPO_SMTXA
39	102	3.8	1645	1	OMPB_RICTY
40	101.5	3.7	1637	1	MRSP_STAUD
41	101	3.7	1012	1	FOX2_NEUDR
42	101	3.7	1012	1	POIG_IBDVO
43	100	3.7	466	1	MMOB_RAT
44	99.5	3.7	807	1	APSK_STRGR
45	99.5	3.7	3567	1	ERY2_SACER

ALIGNMENTS

RESULT	ID	Score	Query Match	Length	ID	Description
1	DHET_ACEBU	483.5	17.9	739	1	DHET_ACEBU
2	DHET_ACEBU	479	17.7	738	1	DHET_ACEBU
3	DHET_ACEBU	477.5	17.7	742	1	DHET_ACEBU
4	DHET_ACEBU	466	17.2	757	1	DHET_GLUOX
5	DHET_ACEBU	435	16.1	600	1	XOXF_PARDE
6	DHET_ACEBU	434	16.0	626	1	DHET_METEX
7	DHET_ACEBU	434	16.0	631	1	DHET_METEX
8	DHET_ACEBU	422.5	15.6	631	1	DHET_PARDE
9	DHET_ACEBU	417.5	15.4	631	1	DHET_PARDE
10	DHET_ACEBU	393	14.5	571	1	DHET_METEX
11	DHET_ACEBU	342	12.6	796	1	DHET_ACEBU
12	DHET_ACEBU	323.5	12.0	801	1	DHET_ACEBU
13	DHET_ACEBU	295	10.9	808	1	DHET_GLUOX
14	DHET_ACEBU	253	9.4	790	1	DHET_ACEBU
15	DHET_ACEBU	231.5	8.6	799	1	DHET_ACEBU
16	DHET_ACEBU	128	4.7	827	1	DHET_ACEBU
17	DHET_ACEBU	126	4.7	799	1	DHET_ACEBU
18	DHET_ACEBU	118.5	4.4	593	1	DHET_ACEBU
19	DHET_ACEBU	116	4.3	353	1	DHET_ACEBU
20	DHET_ACEBU	115.5	4.3	353	1	DHET_ACEBU
21	DHET_ACEBU	115.5	4.3	353	1	DHET_ACEBU
22	DHET_ACEBU	115	4.3	353	1	DHET_ACEBU
23	DHET_ACEBU	115	4.3	353	1	DHET_ACEBU
24	DHET_ACEBU	114.5	4.2	353	1	DHET_ACEBU
25	DHET_ACEBU	114	4.2	353	1	DHET_ACEBU
26	DHET_ACEBU	110.5	4.1	353	1	DHET_ACEBU
27	DHET_ACEBU	110.5	4.1	353	1	DHET_ACEBU
28	DHET_ACEBU	108.5	4.0	353	1	DHET_ACEBU
29	DHET_ACEBU	106.5	3.9	353	1	DHET_ACEBU
30	DHET_ACEBU	106	3.9	353	1	DHET_ACEBU
31	DHET_ACEBU	105	3.9	353	1	DHET_ACEBU
32	DHET_ACEBU	105.5	3.8	353	1	DHET_ACEBU
33	DHET_ACEBU	103.5	3.8	353	1	DHET_ACEBU

Query Match 17.9% ; Score 483.5 ; DB 1 ; Length 739;

Best Local Similarity 26.2%; Pred. No. 1.3e-28;
Matches 153; Conservative 77; Mismatches 237; Indels 117; Gaps 16;

```

QY 10 SAGALLALLAPAPAQVPTVDE-----LLANPAGEMISYQONENYHSPLOTITT 61
DB 17 TAGTICAAISGATAMASADDGATGATGAIHADHPNMNTYGRYSEQRISPLDQINR 76
QY 62 ENNGOOLVWAKMOPGVQV--TPLIHGVMYLANPGVIAQALDAKTDGLIWEHRRLP 119
DB 77 SNGNKLAMVYLDLDTNRQEGPLVIDGVATITNMSMKRAVDAAGKILMSYDPRVPG 136
QY 120 NIA-----TLNS-----FGI----- 129
DB 137 NIADKCCCTVNRGAAYNNGKYTFETFDGLIALDAKTKLWMSVNTIPPEALGKORSY 196
QY 130 -----VANGYVAGSTCOYSPFGC--FVSGHDSATGEELMNYFIIPRAGEGDETCGN 180
DB 197 TYDGAAPRIAKGRVITIGN--GGSEFGARGFYAPFAEIGKVDKMFETAPNPKNEPDTASD 254
QY 181 D-----YEARMTGA-----NGQITDPVTNLVHYGSTAVGPASETORGTGGT 224
DB 255 SVLANKAVQWSPGAMTRGGGCTVDSIVDPAVLVYLGNGSPNNYRFBEGKD 314
QY 225 LYGTRAVAPRDTGELVYMRHQTLPKDMDOECTFEAMVTNVDPQSTMEGLSINPNA 284
DB 315 NLFGSIVALKPTEGVEYVHFQETPMQDMFTSVQOIMTLDPINGET----- 362
QY 285 ATGERRVLTGVPCKTGTMMQFOAETGEFLMADNTYONMIESID-ENGIVTNEADILKE 343
DB 363 -----RHVYHAP--KNGFFIYIDAKTGEFTSGKNYVYVWASGLDKTRIPNDALVTL 417
QY 344 LDVEYDCPTFLGRMPSPALNPDGSIYFIPLNNVCYDMAVDOEFT-SMDVYNTS--- 399
DB 418 TGEKMGIFEDDLGHNFAAMAFSPKGTGLVYIPAOVPLTYNOVGSTPHDSNGLDLM 477
QY 400 NTKLPPG-----KDMIGRIDIDISTGRTLSVRAAANYSPVLSGGVLENGCT 451
DB 478 NKVGIDSPKAOAFKDKLKWIVAMDPQKAEAMRDHKGPMWGGILATGDLPLQGLA 537
QY 452 DRYFRALSGEGETLMQRIATVASGAISTEVDGMOYVIAAG 495
DB 538 NEFHAYDATNGSDLEHFAADSGLIAPVPTVYLANGKQYAVEVG 581

RESULT 2
DHET_ACPEO
AC P28036; STANDARD; PRT; 738 AA.
DT 01-AUG-1992 (Rel. 23, Created)
DT 15-AUG-1992 (Rel. 23, Last sequence update)
DE Alcohol dehydrogenase [acceptor] precursor (EC 1.1.99.8).
OS ADHA.
OC Acetobacter polyoxogenes.
OC Bacteria; Proteobacteria; alpha subdivision; Acetobacteraceae;
OC Acetobacter.
OX NCBI_TaxID=439;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=91159482; PubMed=2001402;
RA Tamaki T., Fukaya M., Takemura H., Tayama K., Okumura H.,
RT Kawamura Y., Nishiyama M., Horinouchi S., Beppu T.;
RT "Cloning and sequencing of the gene cluster encoding two subunits of
RT membrane-bound alcohol dehydrogenase from Acetobacter polyoxogenes.";
RL Biochim. Biophys. Acta 1088:292-300(1991).
CC -1- FUNCTION: DEHYDRATION OF PRIMARY ALCOHOLS (EXCEPTION: METHANOL);
CC -1- CATALYTIC ACTIVITY: A primary alcohol + acceptor = an aldehyde +
CC reduced acceptor.
CC -1- COFACTOR: POO AND HEME.
CC -1- SUBUNIT: HETEROTETRAMER (DEHYDROGENASE, CYTOCHROME, AND TWO
CC SMALLER UNKNOWN SUBUNITS) THAT FORMS THE ALCOHOL DEHYDROGENASE
CC COMPLEX.

```

```

CC CC -1- SUBCELLULAR LOCATION: MEMBRANE-BOUND, FACING THE PERIPLASMIC
CC CC SPACE (POTENTIAL).
CC CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POO DEHYDROGENASE FAMILY.
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (see http://www.isb.ch/announce/
CC CC or send an email to license@isb.sib.ch).
CC DR EMBL: D00635; BAA00528.1;
CC DR PIR: S14270; S14270.
CC DR HSRP: Q924U7; S14270.
CC DR InterPro: IPR001479; Bac_POO.
CC DR InterPro: IPR002372; Bac_POO_repeat.
CC DR Pfam: PF01011; Bacterial_POO.
CC DR PROSITE: PS00363; BACTERIAL_POO_1; 1.
CC DR PROSITE: PS00364; BACTERIAL_POO_2; 1.
CC DR PROSITE: PS00190; CYTOCHROME_C_1.
CC KW Oxidoreductase; POO; Heme; Periplasmic; Membrane; Signal.
CC FT SIGNAL 1 35
CC FT CHAIN 36 738
CC FT BINDING 650 650
CC FT BINDING 653 653
CC FT METAL 654 654
CC FT METAL 654 654
CC SQ SEQUENCE 738 AA; 80840 MW; 1E2B6ED7BCD92AF6 CRC64;

```

Query Match 17.7%; Score 479; DB 1; Length 738;
Best Local Similarity 25.9%; Pred. No. 2.8e-28;
Matches 153; Conservative 77; Mismatches 230; Indels 130; Gaps 16;

```

QY 10 SAGALLALLAPAPAQVPTVDE-----LLANPAGEMISYQONENYHSPLOTITT 61
DB 17 TAGTICAAISGATAMASADDGATGATGAIHADHPNMNTYGRYSEQRISPLDQINR 76
QY 62 ENNGOOLVWAKMOPGVQV--TPLIHGVMYLANPGVIAQALDAKTDGLIWEHRRLP 119
DB 77 SNGNKLAMVYLDLDTNRQEGPLVIDGVATITNMSMKRAVDAAGKILMSYDPRVPG 136
QY 120 NIA-----TLNS-----FGI----- 129
DB 137 NIADKCCCTVNRGAAYNNGKYTFETFDGLIALDAKTKLWMSVNTIPPEALGKORSY 196
QY 130 -----VANGYVAGSTCOYSPFGC--FVSGHDSATGEELMNYFIIPRAGEGDETCGN 180
DB 197 TYDGAAPRIAKGRVITIGN--GGSEFGARGFYAPFAEIGKVDKMFETAPNPKNEPDTASD 254
QY 176 -----ETWGNDEYEAARMYGA-----NGQITDPVTNLVHYGSTAVGPASETOR 218
DB 255 VLMNKAVQWMS-----PGCANTROGGGCTVDSIVDPAVLVYLGNGSPNNYRFBEGKD 314
QY 219 GTPGGLTGTNTRAVAPRDTGELVYMRHQTLPKDMDOECTFEAMVTNVDPQSTMEGLS 278
DB 308 SEGKDNVLFGSIVALKPTEGVEYVHFQETPMQDMFTSVQOIMTLDPINGET----- 362
QY 279 SINPNAATGERRVLTGVPCKTGTMMQFOAETGEFLMADNTYONMIESID-ENGIVTNE 337
DB 358 -----NGETRIYVHARKNGFFIYIDAKTGEFTSGKNYVYVWASGLDKTRIPNDALV 417
QY 338 DATLKLDEYDVDCPTFLGRMPSPALNPDGSIYFIPLNNVCYDMAVDOEFT-SMDVY 396
DB 411 DALYTLTGKEMWGIPODLGHNFAAMAFSPKGTGLVYIPAOVPLTYNOVGSTPHDSN 470
QY 397 NTS--NWKLPPG-----KDMIGRIDIDISTGRTLSVRAAANYSPVLSGGVLENGCT 451
DB 471 NLGLDMNKVIGIPDSPEAKQAFVKDKLKWIVAMDPQKAEAMRDHKGPMWGGILATGDL 530
QY 446 LKNGTDYFRALSGEGETLMQRIATVASGAISTEVDGMOYVIAAG 495

```

Db 531 LFOGLANGERHAYDATNGSDLFHFAADSGIIPVYTLANGKOYAVEVG 580

RESULT 3

DHET_ACEAC STANDARD; PRT; 742 AA.
 AC P18278:
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 01-NOV-1990 (Rel. 16, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Alcohol dehydrogenase [acceptor] precursor (EC 1.1.99.8).
 GN ADHA OR ADHI.
 OS Acetobacter aceti.
 OC Acetobacteriaceae; Proteobacteria; alpha subdivision; Acetobacteraceae;
 OC Acetobacter.
 OC NCBI_TaxID=435;
 RN (1)
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 36-50.
 RA MEDLINE=89255070; PubMed=2722742;
 RA Inoue T., Sunagawa M., Mori A., Imai C., Fukuda M., Takagi M.,
 Yano K.;
 RT Cloning and sequencing of the gene encoding the 72-kilodalton
 RT dehydrogenase subunit of alcohol dehydrogenase from Acetobacter
 RT aceti.
 RL J. Bacteriol. 171:3115-3122(1989).
 RN [2]
 RP 3D-STRUCTURE MODELING.
 RX MEDLINE=95289364; PubMed=7772016;
 RA Gdzier G.E., Giles I.G., Anthony C.;
 RT The structure of the quinoprotein alcohol dehydrogenase of
 RT Acetobacter aceti modelled on that of methanol dehydrogenase from
 RT Methylobacterium extorquens.
 RL Biochem. J. 308:375-379(1995).
 RL -1- CATALYTIC ACTIVITY: A primary alcohol + acceptor = an aldehyde +
 reduced acceptor.
 CC -1- COFACTOR: POQ AND HEME.
 CC -1- SUBUNIT: Tetramer of non identical chains (dehydrogenase,
 CC cytochrome, and two smaller unknown subunits).
 CC -1- SUBCELLULAR LOCATION: Periplasmic.
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POQ DEHYDROGENASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: D90004; BAA14058.1;
 DR EMBL: J50326; J50326.
 DR HSSP: 092437; 1F1G.
 DR InterPro: IPR001479; Bac_POQ.
 DR InterPro: IPR002372; Bac_POQ_repeat.
 DR InterPro: IPR000345; CYC_heme_bind.
 DR Pfam: PF01011; Bacterial_POQ_6.
 DR PROSITE: PS00363; BACTERIAL_POQ_1;
 DR PROSITE: PS00364; BACTERIAL_POQ_2;
 DR PROSITE: PS00190; CYTOCHROME_C_1.
 DR OXidoreductase; POQ; Heme; Periplasmic; Membrane; Signal.
 KM SIGNAL 1 35
 FT CHAIN 36 742 ALCOHOL DEHYDROGENASE [ACCEPTOR].
 FT ACT SITE 343 742 BASE (POTENTIAL) (BY SIMILARITY).
 FT BINDING 649 649 HEME (COVALENT) (BY SIMILARITY).
 FT BINDING 652 652 HEME (COVALENT) (BY SIMILARITY).
 FT METAL 653 653 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
 SEQUENCE 742 AA; 81521 MW; 96G69268DAB825A CRC64;

Query Match 17.7%; Score 477.5; DB 1; Length 742;
 Best Local Similarity 26.2%; Pred. No. 3.7e-28;
 Matches 161; Conservative 89; Mismatches 202; Indels 163; Gaps 24;

QY 11 AGALLAARAFAPVPEVDEL--ANPAGEMISYGOENYRHSPLTOTTEVGOLO 68
 DB 23 AALPYAVAPPARADGCGNTGEIHHADDPENMLSYGRYSQRYSPDQIRNSVNGDK 82
 QY 69 LVYARGMOPGRVQ--VTPLIHDGYVLANPGVYQALIDAKTGDLIMEHRQLP-NIA----- 122
 DB 83 ILGYTTLDTNRGQETPLVVDGIMVATTNMSKMEALDAATGKLIMQYDPKVGNTLADKGC 142
 QY 123 --TLN-----SEGI----- 129
 DB 143 CDTVRKAGYNGKVFCTGTPGRVLVAADAKTGKKVMVNTIPADASIGKORSTYDGAVR 202
 QY 130 VANGVIVAGSTCCQSPFC--FVSGHDSATGEELMRVYFIPRAGEGD----- 175
 DB 203 VAKGVLJLN--GGAFFGARFVSADFDETKLKRFTYVPPNKNEDHAASDNLINRA 260
 QY 176 -ETWGDYERKMT-----GANGITDVPVNLVHYGSTAVGPASETQKPGSTLY--- 226
 DB 261 YKTWGP--KGAWVRGGGSGTWDLSLYDPVSLTY--LAVG-----NSPMNYKRSR 309
 QY 227 --GN-----TFAVRPDTEIWRHQTLPKRWDOECTFEFMMVTVNDVOSTEMEGLOST 280
 DB 310 GIGSNLFLGISTVLKPEETGEYWHFQATPMDQDYTSVOQTITLDMVK----- 358
 QY 281 NPNATGE--RRLTVGPKTGTWMOFDETGELWADPTNQNMLESID-----EN 330
 DB 359 -----GEMRHVTVHAF--KNGFFVLDATGTGELGSKVNYVMANGDPLTGRPMYNDP 411
 QY 331 GIVTVNEDAILKELDYEDVCPTEFGSRMPSAALNPDSGIFIFLNNVY----- 381
 DB 412 GLYLING-----KFWYGI-PGPIGAHNFMAVSPKTHLVYIARHQLPFRYKNQVGF 463
 QY 382 ----DNMAVDDEFISMVYNTSNVTKLPKGDMDIGRIDAIDISTGRTMSVRAAANTSP 437
 DB 464 KPHADSMVWGLDMTKNGIPLDPE--ARFAYIKDLHGLMLADVPKMTVWKIDHKGPWNG 522
 QY 438 VLSTGGVLFNGTDRYFRALSOETGLMOTRLATVASSQATSYEVDMQVYA----- 491
 DB 523 ILATGGDLFQGLANGERHAYDATNGSDLYKFKDAQSGITAPMTYVNGKOYAVEVG 582
 QY 492 ----TAGGVSYSG 502
 DB 583 GIVPISMGGVGTSG 597

RESULT 4

DHET_GLUOX STANDARD; PRT; 757 AA.
 AC 005342;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Alcohol dehydrogenase [acceptor] precursor (EC 1.1.99.8) (G3-ADH
 DE subunit I).
 GN ADHA.
 OS Gluconobacter oxydans (Gluconobacter suboxydans).
 OC Bacteria; Proteobacteria; alpha subdivision; Acetobacteraceae;
 OC Gluconobacter.
 OC NCBI_TaxID=442;
 RN (1)
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 36-50.
 RC STRAIN=IFO 12528;
 RA MEDLINE=97208225; PubMed=9055427;
 RA Kondo K., Horinouchi S.;
 RT Characterization of the genes encoding the three-component membrane-
 RT bound alcohol dehydrogenase from Gluconobacter suboxydans and their
 RT expression in Acetobacter pasteurianus.
 RL Appl. Environ. Microbiol. 63:1131-1138(1997).
 CC -1- FUNCTION: DEHYDRATION OF PRIMARY ALCOHOLS (EXCEPTION: METHANOL)
 CC (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: A primary alcohol + acceptor = an aldehyde +
 reduced acceptor.
 CC -1- COFACTOR: POQ AND HEME.

```

CC -1- SUBUNIT: HETEROTRIMER (DEHYDROGENASE, CYTOCHROME AND PROTEIN
CC ADHS), THAT FORMS THE ALCOHOL DEHYDROGENASE COMPLEX.
CC -1- SUBCELLULAR LOCATION: MEMBRANE-BOUND, FACING THE PERIPLASMIC
CC SPACE (POTENTIAL).
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POO DEHYDROGENASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: D86375; BAA19753.1; -.
DR HSP: Q9247; 1f1g.
DR InterPro: IPR001479; Bac_POO.
DR InterPro: IPR002372; Bac_POO_repeat.
DR Pfam: PF01011; Bacterial_POO_6.
DR PROSITE: PS00363; BACTERIAL_POO_6.
DR PROSITE: PS00364; BACTERIAL_POO_1; 1.
DR PROSITE: PS00190; CYTOCHROME_C; 1.
DR Oxidoreductase; P00; Heme; Periplasmic; Membrane; Signal.
KM SIGNAL 1 34
FT CHAIN 35 757 ALCOHOL DEHYDROGENASE [ACCEPTOR].
FT MOD_RES 35 35 PYRROITONE CARBOXYLIC ACID.
FT ACT_SITE 342 342 BASE (POTENTIAL).
FT BINDING 653 653 HEME (COVALENT) (BY SIMILARITY).
FT BINDING 656 656 HEME (COVALENT) (BY SIMILARITY).
FT METAL 657 657 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
SQ SEQUENCE 757 AA; 82968 MW; 39B9F90E3B947581 CRC64;

Query Match 17.2%; Score 466; DB 1; Length 757;
Best Local Similarity 25.0%; Pred. No. 2.8e-27;
Matches 153; Conservative 89; Mismatches 227; Indels 144; Gaps 19;

OY 7 LMSAGLALIAA--PAPAQV---TPVDELLANPAGEWMSYONCENYHSHPTITTT 61
DB 16 LLSGAAALFAFSAVPAFAQEDTGTATSSDNGHP--GDWLSYGRSTSEORYSPLDQINT 74
OY 62 ENVGOLQVWARGMOPGKQV--TPIIHGVWYLANPGDVIQAIQDAKTGDIWEHRRLP- 119
DB 75 ENVGKTLAMHVDLPINRQESTPLIVGVWYATTNMSKMLDAAGKLLMYDPKVG 134
OY 120 NTA----- 122
DB 135 NIADNGCCDVTYSGAAYWNGKVFETPDGRLLALDAKTGLVWSVYTIPEKQOLGHORSY 194
OY 123 TLNSEGIANGVIYAGSCQYSPGC--FVSGHDSATGEELMHNFTIPRAGEGD----- 175
DB 195 TVDGAAPRIAKKRVILGN--GAEFGANGFVSAPDAETSLDMNFFVVPENKPPGGAASD 252
OY 176 -----ETWGNDEYEARMMTG--AWGQITDPVNTLVHYSGTAVGASERQGTPEGT 224
DB 253 DILMSKAVPTWKGKNGAKWQGGGTVWDSLYDPVTDVLYLVGNGSPNNKFRSEKGD 312
OY 225 LYGTNFEAVAPDGEIWRHQTLPDNDMDCEFTAMVTANVDPSTMEGLASINPNA 284
DB 313 NLFLGSLVALINPDTGKVMHFQETPMDEWDYTSVOOIMTLMPV----- 356
OY 285 ATGE-RNVLTVGVPCKTGMOPDAETGEFLMARDNTYONMIESIDE-NGIYTVNEDAILK 342
DB 357 -NEMRHVIVHAP-KNGFVIYIDAKTGKFIITGKPTTYENMANGLDPRVGRNVPDALMT 414
OY 343 ELDEYDVCTPLGGRDMSALNDSCGTYIPLNNV-----CYDMAVADOE 389
DB 415 LTGKPMVLGIPDELGHNFPAAMAYSPTKLYIYPAQIILYLDQKGGFKAYIHAMNGLD 474
OY 390 FTSMDYNTSNVTKLPKQGMIGRIDIDISTGRL-MSVGRANANS-----PVL 439
DB 475 MNKIGLFDNDDEHVAAKKDF-----LKYVKGWTVAMDEPKMAPAETINHKGPNNGLL 528

```

```

OY 440 STGGVLENGTDRYFRALSGEETLMQTRLATVASGCAISYVDGMQYAI----- 492
DB 529 ATAGNVIVQIANGLEFNAIYATNTGNDLYSPPAOSAIAPVYTYANGKQYAVEVGWGI 588
OY 493 ---AGGVSYSG 502
DB 589 YPFLYGVARTSG 601

RESULT 5
ID XOXF_PARDE STANDARD; PRT; 600 AA.
AC P29968;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Putative dehydrogenase XOXF precursor (EC 1.1.99.-).
GN XOXF.
OS Paracoccus denitrificans.
OC Bacteria; Proteobacteria; alpha subdivision; Rhodobacter group;
OC Paracoccus.
OX NCBI_Taxid:266;
RN [1]
RP SEQUENCE FROM N.A.
RA Harms N.;
RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 380-600 FROM N.A.
RC STRAIN-PD 1235;
RA MEDLINE=92041583; PubMed=1657873;
RA Ras J., Reijnders W.N.M., van Spanning R.J.M., Harms N., Olmann L.F.,
RA Stoutamer A.H.;
RT "Isolation, sequencing, and mutagenesis of the gene encoding
RT cytochrome c553 of Paracoccus denitrificans and characterization of
RT the mutant strain.";
RL J. Bacteriol. 173:6971-6979(1991).
CC -1- COFACTOR: P00 (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POO DEHYDROGENASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: U34346; AAC44555.1; -.
DR EMBL: M75583; AAA25574.1; -.
DR PIR: A41378; A41378.
DR HSP: P38539; 4AAH.
DR InterPro: IPR002372; Bac_POO_repeat.
DR Pfam: PF01011; Bacterial_POO_7.
KM Oxidoreductase; P00; Signal.
FT SIGNAL 1 21
FT CHAIN 22 600 PUTATIVE DEHYDROGENASE XOXF.
FT ACT_SITE 318 318 BASE (POTENTIAL).
SQ SEQUENCE 600 AA; 65159 MW; DCA96F1BCC5A3CE CRC64;

Query Match 16.1%; Score 435; DB 1; Length 600;
Best Local Similarity 25.7%; Pred. No. 4.3e-25;
Matches 154; Conservative 80; Mismatches 227; Indels 138; Gaps 23;

OY 10 SAGALALIAAPAPAQVPTVDELLANP-----PAGEWISYONCENYHSHPTITTE 62
DB 6 NGACTALLMSGTA-----LANEQRAGRDQAPQMAIQMGYANTRYSTLDQINKD 56
OY 63 NVGQLOLVW--ANGMOPGVQVYPTLTHGVWYLANP-GDVIQAIQDAK-TGDLTWEHR-RQ 117
DB 57 NVKDLRVAMTFSTGVLRGH-EGSPVLYIGVWYVHPPFNRRFALDLNDNGKILMYRPEQO 115

```

```

QY 118 LPIATLNSFGIYANGVYAGSTCOYSPGCGFSGHDSATGEELMRNF-IPRAGE----- 172
RT 118 LPIATLNSFGIYANGVYAGSTCOYSPGCGFSGHDSATGEELMRNF-IPRAGE----- 172
RT 116 DPNVIAWMSDFYRGLSYADGMILLGQADTTVALDITSEYVMSKIKIDPGIGELTA 175
DB 116 DPNVIAWMSDFYRGLSYADGMILLGQADTTVALDITSEYVMSKIKIDPGIGELTA 175
QY 173 -----EGDET-----GNDYEA----- 184
DB 176 TVPVPKRVLVGISGGEYVGRMTALNLTDSSEAMKMSSTGPDELLVDETTILGKP 235
QY 185 -----FMTGA-----WGQITTPYVNLVHGSTAVGAPASEGTGPGTLYGT 228
DB 236 IGADSSINSWEGOWOIGGGTITGWFSTYDPLNLVYGTGNPSTWNPQR-PQDNKW-S 292
QY 229 NTRAVAPDGEIYVHRLPRDNMDOCTFEKMTN-VDPQSTMEGLQSTINPNAAT 286
DB 229 NTRAVAPDGEIYVHRLPRDNMDOCTFEKMTN-VDPQSTMEGLQSTINPNAAT 286
DB 293 MTIMARDADGMAMKRYOMTPHDEMVDYCNEMILLNQTVDGQ----- 335
QY 287 GERVLTVGPKCTGMQFDETEGEELMARDNYQMAIE-----SIDENGIYVEDAIL 341
DB 336 -ERKLTTFHD-RNGLATYMDRETGELLVAE-RYDPVVMWTTGVDDPNS-ETVGRPAV 390
QY 342 KELDEYD-----VCPTEFGGDMPSAALNPDGIYFIPLNNVCYDMAVDOETSM 393
DB 391 AEYTAONGEDETNTGCPALATKIDQPAFSPKTNLFVPTNHVCMDEEPFVAYTAG 450
QY 394 DYY-NTSNVTKLPKGMIDRIDALDITGRTLSYERAAVSTGSGVLFNGGT 451
DB 451 QYVYATLSMTPAPNSHCGMNFAMHNTGELKMSVPEQSVWSGALTAGVYFYTGL 510
QY 452 DRYFRLAISOGEITLMOTRLATVASGQAISEYVDGMQVYALAG-----GVSYGSLNS 505
DB 511 EGYLKRVDAQTEELKRTKPTSGIIGNVMTYEHGKRYGVLISGVGMAGIAGLTLN 569

RESULT 6
DHML_METEX STANDARD: PRT: 626 AA.
ID DHML_METEX STANDARD: PRT: 626 AA.
AC P16027;
DT 01-APR-1990 (rel. 14, Created)
DT 01-APR-1990 (rel. 14, Last sequence update)
DT 15-JUL-1999 (rel. 38, Last annotation update)
DE Methanol dehydrogenase subunit 1 precursor (EC 1.1.99.8) (MDH large
alpha subunit) (MEDH).
GN MOXF.
OS Methylobacterium extorquens.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Methylobacterium group; Methylobacterium.
OX NCBI_TaxID=408;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AM1 / NCIB 9133;
RC MEDLINE=90337342; PubMed=2116368;
RA Anderson D.V., Morris C.J., Nunn D.N., Anthony C., Lidstrom M.E.;
RT "Nucleotide sequence of the methylobacterium extorquens AM1 moxf and
RT moxf genes involved in methanol oxidation."
RT Gene 90:173-176(1990).
RN [2]
RP SEQUENCE OF 28-53.
RC STRAIN=AM1 / NCIB 9133;
RC MEDLINE=89350892; PubMed=2504152;
RA Nunn D.N., Day D., Anthony C.;
RT "The second subunit of methanol dehydrogenase of methylobacterium
RT extorquens AM1."
RT Biochem. J. 260:857-862(1989).
RN [3]
RP DISULFIDE BONDS.
RC MEDLINE=95584759; PubMed=7656012;
RA Blake C.C.F., Ghosh M., Harlos K., Avezoux A., Anthony C.;
RT "The active site of methanol dehydrogenase contains a disulphide
RT bridge between adjacent cysteine residues."
RT Nat. Struct. Biol. 1:102-105(1994).
RN [4]
RP X-RAY CRYSTALLOGRAPHY (1.94 ANGSTROMS).
RC MEDLINE=95253818; PubMed=7735834;

```

```

RA Ghosh M., Anthony C., Harlos K., Goodwin M.G., Blake C.;
RT "The refined structure of the quinoprotein methanol dehydrogenase
RT from methylobacterium extorquens at 1.94 A."
RT Structure 3:177-187(1995).
RL 1-1 CAPALYTIC ACTIVITY: A primary alcohol + acceptor - an aldehyde +
CC reduced acceptor.
CC -1- COFACTOR: PQO.
CC -1- SUBUNIT: THE HOLOENZYME MDH HAS AN ALPHA-2/BETA-2 CONFIGURATION
CC OF A LARGE ALPHA AND A SMALL BETA SUBUNIT.
CC -1- SUBCELLULAR LOCATION: PERIPLASMIC, BUT ASSOCIATED WITH INNER
CC MEMBRANE.
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL PQO DEHYDROGENASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: M31108; AAA25380.1; -.
CC PIR: S07908; S07908.
CC PIR: J00706; J00706.
CC HSP: P38539; AAAH.
CC InterPro: IPR001479; Bac_PQO.
CC InterPro: IPR002372; Bac_PQO_repeat.
CC Pfam: PF01011; Bacterial_PQO_7.
CC PROSITE: PS00363; BACTERIAL_PQO_1; 1.
CC PROSITE: PS00364; BACTERIAL_PQO_2; 1.
CC Oxidoreductase; PQO; Signal; Methanol utilization; Periplasmic.
CC SIGNAL 28 27
CC CHAIN 1 626 METHANOL DEHYDROGENASE SUBUNIT 1.
CC FT DISULFID 130 131
CC FT DISULFID 413 442
CC FT ACT_SITE 330 330
CC FT ACT_SITE 68434 MM; 649880DAPD2AD34C CRC64;
CC SEQUENCE 626 AA; 68434 MM; 649880DAPD2AD34C CRC64;

Query Match 16.0%; Score 434; DB 1; Length 626;
Best Local Similarity 24.5%; Pred. No. 5,4e-25;
Matches 148; Conservative 91; Mismatches 208; Indels 158; Gaps 23;

QY 10 SAGALALLA-AP-AFOVYPTVDELANPPAGE-WISYONGONENRHSPLDTITENYQ 66
DB 7 SVSALAMALAPALSSGAYANDKLVELSKSDNWMVPGKNDSDNNSFSLKINKGNKQ 66
QY 67 IQLVV-ARGMOPKRYOYTPLIHDGVV-----LANPQV----- 99
DB 67 LRPATFSTGLNGH-EGAPLVYDGKMYHTSPNNTFALGLDDPGLIMODKPKONPAA 125
QY 100 -----IOAKTGDLLMEHRRQLPNATLN 125
DB 126 RAVACODLVNRLGAWPDPGTPALILKTLQDGNVALNMEETETVWK----- 173
QY 126 SFGIYANGVYAGSTCOYSPF-----GC-----FVSGHDSANGEEELMRNF- 167
DB 174 ----VENSIDIKVSTLITLAFYVVDKVTIISGAGELGVRGLYADVYGEQVYRAYATG 229
QY 168 -----PRAGEE--GDETWGNDYEARMWGA-----WGQITTPYVNLVHYS 207
DB 230 PDKDLLASDFNIKPNHYGOKGLGTGTEGD---AWKIGGCTNGWYAYDPGNTLIFG 286
QY 208 TAVGPASEFQRTPGTILGNTTPRAVVRPDGEIYVHRLPRDNMDOCTFEKMTNVD 267
DB 287 GNPAPNMTWR--PGDNKM-TWTIFGRADYGEAKFYQKTPHDEMD-----YAGVN 335
QY 268 VDPSTMEGLQSTINPNAATGERVRLTVGPKCTGMQFDETEGEELMRN- -DTNYQNM 324
DB 336 VMMSSEKQD-----KDGAKRKLTHPDNRGIVYTLTDRLDGLVYSANKLDDF--VNVF 385
QY 325 ESIDENGIYVNEDELAKELD--VEYDVCTFLGGRDWSAALNPDGSIYFIPLNNVCYDM 363

```


Db 386 KSVDLKTGQPVDPPEYGRMDHLAKDIPSAHNGHSDYPPKKELEFFMGINHICMDW 445
 QY 384 MAVDEFTSMQVYNTSNTKLPCK-----DMIGRIDALIDISTGRTLSVERAAANYSP 437
 Db 446 EPEMLPYRAGQFFVGATLMNTPGPKGRONTEGIGQIKAYNAITGDYKKEKMERFAVWG 505
 QY 438 VLSGGVLFNGGTRDFRFRALSOETGELLWOTRLATVAASGAATSYEDVMQVVAIAGGV 497
 Db 506 TMAATAGDLVFGTLDYLKARDSDTGLMKFKIPSGAIGYPMYTHKGTGYAL----- 560
 QY 498 SYGSG 502
 Db 561 YYGVG 565

RESULT 7
 DHM1_METOR
 ID DHM1_METOR STANDARD: PRT: 626 AA.
 AC P15279;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-APR-1990 (Rel. 14, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Methanol dehydrogenase subunit 1 precursor (EC 1.1.99.8) (MDH large
 alpha subunit) (MEDH).
 GN MOXF.
 OS Methylobacterium organophilum XX.
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Methylobacterium group; Methylobacterium.
 OX NCBI_TaxID=410;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 29-43.
 RC STRAIN=ATCC 27886 / DSM 760 / NCIB 11278;
 RX MEDLINE=8908094; PubMed=2459109;
 RA Machlin S.M., Hanson R.S.;
 RT "Nucleotide sequence and transcriptional start site of the
 methylobacterium organophilum XX methanol dehydrogenase structural
 gene".
 RT Methylobacterium organophilum XX methanol dehydrogenase structural
 gene".
 RL J. Bacteriol. 170:4739-4747(1988).
 CC -1- CATALYTIC ACTIVITY: A primary alcohol + acceptor = an aldehyde +
 reduced acceptor.
 CC -1- COFACTOR: POQ.
 CC -1- SUBUNIT: THE HOLOENZYME MDH HAS AN ALPHA-2/BETA-2 CONFIGURATION
 OF A LARGE ALPHA AND A SMALL BETA SUBUNIT.
 CC -1- SUBCELLULAR LOCATION: PERIPLASMIC, BUT ASSOCIATED WITH INNER
 MEMBRANE.
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POQ DEHYDROGENASE FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@sib-sib.ch).
 CC EMBL: M22629; AAS0289.1;
 DR HSSP: P38539; 4AAH.
 DR InterPro: IPR001479; Bac_POQ.
 DR InterPro: IPR002372; Bac_POQ_repeat.
 DR Pfam: PF01011; Bacterial_POQ_7.
 DR PROSITE: PS00363; BACTERIAL_POQ_1; 1.
 DR PROSITE: PS00364; BACTERIAL_POQ_2; 1.
 KW Oxidoreductase; POQ; Signal; Methanol utilization; Periplasmic.
 FT CHAIN 1 28
 FT DISULFID 29 626 METHANOL DEHYDROGENASE SUBUNIT 1.
 FT DISULFID 130 131 BY SIMILARITY.
 FT DISULFID 413 442 BY SIMILARITY.
 FT ACCT_SITE 330 330 BASE (POTENTIAL).
 SO SEQUENCE 626 AA; 68677 MW; 8768PB68371E5DF CRC64;

Query Match 16.0%; Score 434; DB 1; Length 626;
 Best Local Similarity 24.5%; Pred. No. 5.4e-25;

Matches 148; Conservative 91; Mismatches 208; Indels 158; Gaps 23;
 QY 10 SAGALALLA-AP-APAOYTPYDELALNPAGE-MISYGOQEVYRSPITQITENVOQ 66
 Db 7 SVSALMLALAPALSSVAYANDKVELSKDDVMWPKNTSDNNSELKOVNKSVMQ 66
 QY 67 LQLVW--ARGNPGKVQVYTPPLIHGVMY-----LANPGDV----- 99
 Db 67 LRPATFTSTGLNGH-ESAPLVVDKMYVHTSPFNNTFALDLDDPGHLLMQKPKQNPA 125
 QY 100 -----IOADKATGDLTMEHRROLPIATLN 125
 Db 126 RAVACCDLVNRLAWPQDGKTPALLIKTQDLDRHYVALNKEIGETVWK----- 173
 QY 126 SFGIVANGVIVAGSTCOQSPF-----GC-----FVSGHDSATGELMARNFY- 167
 Db 174 -----VENSIDIKVSGTLTIAPYVKKVITIGSSGAELGVGYLAVDYKGGVWRATNG 229
 QY 168 -----PRAGEE--GDETWNDEYFARMNKA-----KQITYPYTNLVHYS 207
 Db 230 PDKDLLADDPENYKNAHYGKGIGTATWBGD--AKKIGGTNMGWVAYIDGTNLIYFGT 286
 QY 208 TAVGPASETORGPFGGLYGTNTREAVRPDGEIYWRHQITLPRMDDOCTEFEMAVTND 267
 Db 287 GNPAFNMETWR--PGDKW-TMTIFGRDADTEAKFGYOKPHDEMDYAGVNMW----- 338
 QY 268 VQSTEMEGLOSTINPNAATGERVLTGVCPTGMQFDATGEPLMAR--DTNYQNT 324
 Db 339 --PSEKQD-----KDKTRKLLTHEDRNGIYTLTDRTDGLAVSNKLDFT--VNVF 385
 QY 325 ESIDENKIVYVNDALIKELD-VEVDVCPTELGGRDPSAALNDSGIYFPLNNVYDM 383
 Db 386 KTYVDLTGQVRPPEYGRMDHLAKDIPSAHNGHSDYPPKKELEFFMGINHICMDW 445
 QY 384 MAVDEFTSMQVYNTSNTKLPCK-----DMIGRIDALIDISTGRTLSVERAAANYSP 437
 Db 446 EPEMLPYRAGQFFVGATLMNTPGPKGRONTEGIGQIKAYNAITGDYKKEKMERFAVWG 505
 QY 438 VLSGGVLFNGGTRDFRFRALSOETGELLWOTRLATVAASGAATSYEDVMQVVAIAGGV 497
 Db 506 TMAATAGDLVFGTLDYLKARDSDTGLMKFKIPSGAIGYPMYTHKGTGYAL----- 560
 QY 498 SYGSG 502
 Db 561 YYGVG 565

RESULT 8
 DHM1_PARDE
 ID DHM1_PARDE STANDARD: PRT: 631 AA.
 AC P12293;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Methanol dehydrogenase subunit 1 precursor (EC 1.1.99.8) (MDH large
 alpha subunit) (MEDH).
 GN MOXF.
 OS Paracoccus denitrificans.
 OC Bacteria; Proteobacteria; alpha subdivision; Rhodobacter group;
 OC Paracoccus.
 OX NCBI_TaxID=266;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 33-49.
 RC MEDLINE=87307969; PubMed=3114231;
 RA Harms N., de Vries G.E., Maurer K., Hoogendijk J., Stoutamer A.H.;
 RT "Isolation and nucleotide sequence of the methanol dehydrogenase
 structural gene from Paracoccus denitrificans".
 RL J. Bacteriol. 169:3969-3975(1987).
 CC -1- CATALYTIC ACTIVITY: A primary alcohol + acceptor = an aldehyde +
 reduced acceptor.
 CC -1- COFACTOR: POQ.
 CC -1- SUBCELLULAR LOCATION: Periplasmic.
 CC -1- MISCELLANEOUS: MDH IS THE MAJOR PROTEIN IN THE CELL DURING GROWTH

ON METHANOL (IN P-DENITRIFICANS MDH CONSTITUTES UP TO 15% OF THE TOTAL CELL PROTEIN).
 -1- SIMILARITY: BELONGS TO THE BACTERIAL POQ DEHYDROGENASE FAMILY.
 THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC EMBL: M17339; AAA88366.1; .
 CC HSSP: P38539; AAAH.
 CC InterPro: IPR001479; Bac_POQ.
 CC InterPro: IPR002372; Bac_POQ_repeat.
 CC Pfam: PF01011; Bacterial_POQ_7.
 CC PROSITE: PS00363; BACTERIAL_POQ_1; 1.
 CC PROSITE: PS00364; BACTERIAL_POQ_2; 1.
 CC Oxidoreductase; POQ; Signal; Methanol utilization; Periplasmic.
 CC FT SIGNAL 1 32 METHANOL DEHYDROGENASE SUBUNIT 1.
 CC FT CHAIN 33 631 BY SIMILARITY.
 CC FT DISULFID 135 136 BY SIMILARITY.
 CC FT DISULFID 418 447 BASE (POTENTIAL).
 CC FT ACT_SITE 335 335
 CC SEQUENCE 631 AA; 69799 MW; 0934DC93FFC5730B CRC64;
 Query Match 15.6%; Score 422.5; DB 1; Length 631;
 Best Local Similarity 24.8%; Pred. No. 4e-24; Indels 179; Gaps 26;
 Matches 153; Conservative 85; Mismatches 201;
 QY 4 TSLMASAGLALL-AAPAAQVPTVDL--LANPAGWISYGCNENTRSHPLQIT 60
 DB 12 TSLMAAAMGALVITTPARA-----NDQLVELAKDPA-NMVMGTRDYNAONTSEMDIN 65
 QY 61 TENYGOLOLV--ARGM-----OPKV--QVTP- 84
 DB 66 KENKQALPAMSFSTGVLMHGEGTPVYVGRMEIHPPTNTFFALDINPEKILLMKNPK 125
 QY 85 -----LHDCVMTLANPD-----VIOIDAKTGDLMEHRQL 118
 DB 126 ONPRTAVACDVNRGLATV--PGDQVPLFRTOLDGHIYAMDAGETRM- 177
 QY 119 PIATLANSFGIVANGVYASTCOYSPF-----GC-----FSGHDSATGEEL 161
 DB 178 -----IMNSDIKVGSTLTITAPVYIKDLVYSSGAELGVYGYTAVDKSGEMR 227
 QY 162 WRNYFI-----PRAGEE--GDETWNQDYARWMTGA--WQITDPVT 200
 DB 228 WRKFAATGPDELLLAEDFNAPNPHYGKNLIGLETWEGD--AMKIGGTWNGWYAYDPVY 284
 QY 201 NLVHGSTAVPASSTORTPGTGLTGNTPFAVRPRPGETVWHQTLPRDNNDQECTEE 260
 DB 285 DLFYSSGNPAPWNETMR--PEDNKM--TMAIWRGATTEGKAFYOKTPHDEWD----- 335
 QY 261 MANTVNDVOSTEMEGLOSINPNAATGERVLTGVYCKTGWQCPDAETGEFLMARPVY 320
 DB 336 --YAGVAVMALSRQEDKO-----GQWRKLLTHPRDNGIVYLLDRTNGLISA----- 380
 QY 321 QNMIESIDE--NCIVYVNDALIKELDEY-----DVCPTFLGRDPSALNPD 369
 DB 381 -----DKMDTVNWKVEQVOLDTGLPRVDPREGTRMDHKARDICPSAMGHNQSHSDYPER 436
 QY 370 GIVFIPNNVCYDMAVDEFTSMDYNTSNVTKLPEGKDM-----IGRIDAISTGRT 424
 DB 437 KVFMLGINHICDMERPLFYRAGOFFVAGATILMYGPRFATRAERAGACQIKAYDAISEM 496
 QY 425 LMSVRAAANTSVLSTGGVYLRNGGDRYFRALSOETGETLMQFRLAVASGAQISYEV 484
 DB 497 KNEKKERSVWGTMATAGGLTFYVYLDGFIKARDSDTGLMLKFLPSGVIGHPMFTYKH 556
 QY 485 DGMQVVALAGGVSYSG 502

DB 557 DGRQVAL-----MYGVG 569
 RESULT 9
 EXAA_PSEAE STANDARD; PRT; 623 AA.
 AC 0924J77
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Quinoprotein ethanol dehydrogenase precursor (EC 1.1.99.-) (OEDH).
 GN EXAA OR PA1982.
 OS Pseudomonas aeruginosa.
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 OC Pseudomonas
 OX NCBI_TaxID=287;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN-ATCC 17933;
 RX MEDLINE=99041560; PubMed=9826187;
 RA Diehl A., Wintzingerode F., Goerisch H.,
 RT "Quinoprotein ethanol dehydrogenase of Pseudomonas aeruginosa is a
 RT homodimer; sequence of the gene and deduced structural properties of
 RT the enzyme."
 RL Eur. J. Biochem. 257:409-419(1998).
 RN (2)
 RP SEQUENCE FROM N.A.
 RC STRAIN-ATCC 15692 / PA01;
 RX MEDLINE=20437337; PubMed=10984043;
 RA Stover K.K., Pham X.-O.T., Erwin A.L., Mizoguchi S.D., Warren P.,
 RA Hickey M.J., Brinkman F.S.L., Huftagle W.O., Kowalik D.J., Lagrou M.,
 RA Gader R.L., Goltzy L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
 RA Brody J.L., Coulter S.N., Folger K.R., Kas A., Latbig K., Lim R.M.,
 RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
 RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.,
 RA "Complete genome sequence of Pseudomonas aeruginosa PA01, an
 RT opportunistic pathogen."
 RL Nature 406:959-964(2000).
 RN (3)
 RP SEQUENCE OF 1-52 FROM N.A.
 RC STRAIN-ATCC 17933;
 RX MEDLINE=99173751; PubMed=10075429;
 RA Schober M., Goerisch H.,
 RT "Cytochrome c550 is an essential component of the quinoprotein ethanol
 RT oxidation system in Pseudomonas aeruginosa: cloning and sequencing of
 RT the genes encoding cytochrome c550 and an adjacent acetaldehyde
 RT dehydrogenase."
 RL Microbiology 145:471-481(1999).
 RN (4)
 RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF 35-616.
 RC STRAIN-ATCC 17933;
 RX MEDLINE=20202376; PubMed=10736230;
 RA Kettel T., Diehl A., Knaute T., Stezowski J.J., Hoehne W.,
 RA Goerisch H.,
 RT "X-ray structure of the quinoprotein ethanol dehydrogenase from
 RT Pseudomonas aeruginosa: basis of substrate specificity."
 RL J. Mol. Biol. 297:961-974(2000).
 CC -1- FUNCTION: OXIDIZE PRIMARY ALCOHOLS AND ALSO ACTS ON SECONDARY
 CC ALCOHOL, BUT NOT HIGHLY ACTIVE ON METHANOL.
 CC -1- COFACTOR: POQ AND CALCIUM.
 CC -1- SUBUNIT: HOMODIMER.
 CC -1- SUBCELLULAR LOCATION: Periplasmic.
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POQ DEHYDROGENASE FAMILY.
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: AJ009858; CAA0896.1; -
 DR EMBL: AE004624; AAC05370.1; -
 DR EMBL: AF068264; AAC79657.1; -
 DR PDB: 1FLG; 30-AUG-00.
 DR InterPro: IPR001479; Bac_POO.
 DR InterPro: IPR002372; Bac_POO_repeat.
 DR Pfam: PF01011; Bacterial_POO_7.
 DR PROSITE: PS00363; BACTERIAL_POO_1; FALSE_NEG.
 DR PROSITE: PS00364; BACTERIAL_POO_2; FALSE_NEG.
 DR Oxidoreductase; POO; Periplasmic; signal; Calcium; 3D-structure;
 KW Complete proteome.
 FT SIGNAL 1 34
 FT CHAIN 35 623
 FT DISULFID 139 140
 SQ SEQUENCE 623 AA; 68123 MW; 32DDE5DF20B291D6 CRC64;

Query Match 15.4%; Score 417.5; DB 1; Length 623;
 Best Local Similarity 23.2%; Pred. No. 9.2e-24;
 Matches 139; Conservative 94; Mismatches 248; Indels 119; Gaps 19;

OY 3 PSLMASAGALAL--LAAPAPQVPTDELLAN--PRAGEMISYGOQENYRHSPLT 57
 DB 9 PAGLRPSLHCLAEVALAGSAGAILANDKTDGTVLQYKMGTHAQRWSPK 68
 OY 58 QITTEWGOLOLVNARGM--OPKYOVTPLHDGMYTLANGDVIQAIIDAKTGDLWE 113
 DB 69 QVADNVKFLTPAMSYSGDEKQKQ--ESQAIYSDGVIYVTSYSTLFDLAKTGKRLWT 127
 OY 114 HRRLP--N-----IATLNSFGIVAN--GVV-----AGST 140
 DB 128 YNHRPLPDIPRCDCVNRGAIVGDKVFETLDASVALNKNMGVKKKRAHDGAGT 187
 OY 141 COYSP-----FGCF--VSGHDSATGEELRNYPTR----- 169
 DB 188 MTGAPTIKDKTGKVLILHSGSGDEFVGVGRFAPRDPYTGEEIMRPFEGHMGRLNGK 247
 OY 170 -----AGEGSETWGNDEA-----RMWTGA--WGQITVDPVTLVYSGTAQPSSET 216
 DB 248 DSYVTGCVKAPWPDNRNSPTGKVESMSHGGAQWQSAFPAERTITIVGANGPWNWTW 307
 OY 217 QRGTPGG-----TLGTNTREAVRPDGETVYRQTLPRPNMDQECTFEKMTNVDVOP 270
 DB 308 AFAKAGNPHDLSLT--TSGQYGVDPSSGEVKEWFOHTPNDAMFOSNNELVLFYKAKD 366
 OY 271 STEMEGLQGINN-----AATGERVLTGVPCKTGTW--QFDEATGEFIMARDIYQCN 323
 DB 367 GKIVKATAHADNRNGFFIYVDRSGKLONAFPFVDNITWASHIDUKT-----RVEREQ 421
 OY 324 IESIDENGIVVNEDAILKEIDVEDYDVCPTFLGRDPSALNPDGSIYFIPLNNVCYDM 383
 DB 422 RPLPEPG-----QKHGKAVEVSPPLGKNNPMAYSDQDGLPYVPANHWKEDY 471
 OY 384 MAVDQETSMQVYNTSVNTKLPKGMIGRIDAIDISTGRTLSVRAAANYSPILSTGG 443
 DB 472 WTEEVSYTKGSAVILGMGERIKRMVDHVGSLRAMDPVSGVYVWEHKEHLPLWAGVLAAG 531
 OY 444 GVLNNGTDRTFRALSOETGTLMOPTLATVASGQAISEVDGMOVALAGGVSGL 503
 DB 532 NLVFTGTDGGEKFAFKDAKSKELMKFQTGSGIVSPITWEDQGEQYLGVT--VGYGAV 588
 RESULT 10
 DHML_METME
 ID DHML_METME STANDARD; PRT; 571 AA.
 AC P38539;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Methanol dehydrogenase subunit 1 (Ec 1.1.99.8) (MDH large alpha subunit) (MDH).
 OS Methylophilus methylotrophus (Bacterium W3A1).
 OC Bacteria; Proteobacteria; beta subdivision; Methylophilus group;

OC Methylophilus.
 OX NCBI_TaxId=17;
 RN [1]
 RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS).
 RX MEDLINE=9405969; PubMed=8241148;
 RA White S., Boyd G., Mathews F.S., Xia Z.-X., Dai W.-W., Zhang Y.-F.,
 RA Davidson V.L.;
 RT "The active site structure of the calcium-containing quinoprotein
 RT methanol dehydrogenase.";
 RL Biochemistry 32:12955-12958(1993).
 RN [2]
 RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS).
 RX MEDLINE=93054513; PubMed=1331050;
 RA Xia Z.-X., Dai W.-W., Xiong J.-P., Hao Z.-P., Davidson V.L., White S.,
 RA Mathews F.S.;
 RT "The three-dimensional structures of methanol dehydrogenase from two
 RT methylophilic bacteria at 2.6-A resolution.";
 RL J. Biol. Chem. 267:22289-22297(1992).
 CC -1- CATALYTIC ACTIVITY: A primary alcohol + acceptor -> an aldehyde +
 CC reduced acceptor.
 CC -1- COFACTOR: TWO MOLECULES OF POO AND TWO MOLECULES OF CALCIUM
 CC PER TETRAMER.
 CC -1- SUBUNIT: THE HOLOENZYME MDH HAS AN ALPHA-2/BETA-2 CONFIGURATION
 CC OF A LARGE ALPHA AND A SMALL BETA SUBUNIT.
 CC -1- SUBCELLULAR LOCATION: PERIPLASMIC, BUT ASSOCIATED WITH INNER
 CC MEMBRANE.
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POO DEHYDROGENASE FAMILY.
 DR PDB; 4AAH; 08-DEC-96.
 DR InterPro: IPR001479; Bac_POO.
 DR InterPro: IPR002372; Bac_POO_repeat.
 DR Pfam: PF01011; Bacterial_POO_7.
 DR PROSITE: PS00363; BACTERIAL_POO_1; FALSE_NEG.
 DR PROSITE: PS00364; BACTERIAL_POO_2; 1.
 KW Oxidoreductase; POO; Methanol utilization; Periplasmic; 3D-structure;
 KW Calcium.
 FT DISULFID 103 104
 FT DISULFID 379 408
 FT ACT SITE 297 297
 FT HELIX 2 9
 FT TURN 11 12
 FT STRAND 14 14
 FT TURN 17 18
 FT TURN 21 22
 FT STRAND 26 27
 FT TURN 34 36
 FT HELIX 37 39
 FT STRAND 41 47
 FT STRAND 59 61
 FT TURN 62 63
 FT STRAND 64 68
 FT TURN 71 73
 FT STRAND 75 79
 FT TURN 80 81
 FT STRAND 83 84
 FT STRAND 86 90
 FT HELIX 96 101
 FT TURN 103 104
 FT STRAND 112 114
 FT TURN 115 116
 FT STRAND 117 121
 FT TURN 123 124
 FT STRAND 126 131
 FT TURN 132 134
 FT STRAND 137 142
 FT HELIX 146 148
 FT TURN 149 149
 FT STRAND 151 151
 FT STRAND 157 159
 FT TURN 160 161
 FT STRAND 162 166
 FT STRAND 168 168
 FT HELIX 170 172
 FT TURN 173 173
 BASE (POTENTIAL).

[illegible]

Query Match	12.6%	Score 342	DB 1	Length 796
Best Local	Similarity 22.7%	Pred. No. 5	9e-18	
Matches 160	Conservative 74	Mismatches 202	Indels 270	Gaps 32
QY	6 LMAS-----	AGALMLAPFAQVTPVYDDELLANPAGWEISVIGONQENYHSHSTQ		
DB	135 LTMAGNDNPQEIINGTISADTAPRA-ISPVDQ-----	DMPAYGNQSGQSFSPLO		
QY	59 ITTEVNGOLOLWMA-----	RGMPGV--QYTPILHDGVWYLAANGDVYQALDIAKTG		
DB	186 IADVDINLEKAVVFTGDKVQKOPNDPEITNEVTPKAGDLYLCTAHQRILRLDMAISGK			
QY	110 LIWEHRQLPNIATLNSF-----	GI-----		
DB	246 EKNHYD--PELKTNEFHQVTCRGVSYHEAKETAETASPEVMADCCPRILILEVNDGRILAI			
QY	138 ----GSTQ-----	YSP-----		
DB	303 MAENKCLTEFPAKGVLYLNQSNMPTDKILETPITLITTKTIYVAGSTVDNSTRSTS			
QY	148 CFVSGHDSATGELIRNFTIPRA-----	GEGETNGNDYEARMWTGAKQITVDEVT		
DB	363 GVTRGPDVMTGELMA--FDGAKDNPALPSDEHTFFENS-----	PSMAPAYDAKL		
QY	201 NLVHYGSTAVGASGTQRTGGTLYGNTFPANRPDTGEIVYRHQITPRIMDQECTRE			
DB	414 DLYVTPMVTTPDIKMGKRTPEOERYASSI-LALNATTKGLAWSYQIYHDLMDND--			
QY	261 MKNVNVDPSTMEGLQSNPNPNAETERRVLTGVCKTGTGMOFDEGEEL-----			
DB	469 ----LPAOPT-----	LADITVN--GQKPYIYAPAKTGINIFVLRBRNGELVYAPAEKP		
QY	314 ----WADTQYQMKLISIDENGI-----	VTVNDALIKEL--DVEYD-----		
DB	516 VPQAGAKGDTYTPLOPPESELSEFPRTKDSGMDMGATMFDOLVCVMMHQRKEGIFTP			
QY	350 ----VCPFLGLGRMPPSALNP-----	VTNEDALIKEL--DVEYD-----		
DB	576 SEQGLVFEPSGLMEFMGIVSDPREVALANMALPFSVKLIFRGCKNMDQPIDAKGT			
QY	368 ----DSGI-----	YFPLNNCYDMANVDOETSDMYVNTSNTKLPPECKIDIGRIAD		
DB	636 GTESGILQIPQYPPGVATLN-----	PFLS-----		
QY	419 ISYGRILMSYERAAANYS-----	PV-----		
DB	677 LKTNVYMKKRITPODSSHPFMPVPYFPMGNMPLGIGFISACNATLFLMAADVYLAAY			
QY	459 SQETGELIMOTRLATVAVSQAQ--ISYENQKQYVAI-AGGVSQVS			
DB	737 NMSGKELQGNLP--AGQATPMTELEVNGKQYVVISAGHGSFGT			
RESULT 12				
ID	DRGA_ACTICA	STANDARD;	PRT;	801 AA.
AC	P05465;			
DT	01-NOV-1988 (Rel. 09, Created)			
DT	16-OCT-2001 (Rel. 09, Last sequence update)			
DE	Glucose dehydrogenase-A (pyruvate quinolone-glucose DH) (GdhA).			
GN	(ec 1.1.99.17) (Quinoprotein glucose DH) (GdhA).			
OS	Acetibacter calcoaceticus			
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Betaproteobacteria; Rhodospirillaceae; Rhodospirillum rubrum			

CC Acinetobacter.
 OX NCBI_TaxID=471;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=LM2 79.41;
 RX MEDLINE=88289368; PubMed=3399393;
 RA Cleton-Jansen A.-M., Goosen N., Ode G., van de Putte P.;
 RT "Nucleotide sequence of the gene coding for quinoprotein glucose
 dehydrogenase from Acinetobacter calcoaceticus";
 RL Nucleic Acids Res. 16:6228-6228(1988).
 CC -1- CATALYTIC ACTIVITY: D-glucose + acceptor = D-glucono-1,5-lactone +
 reduced acceptor.
 CC -1- COFACTOR: PQQ.
 CC -1- SUBUNIT: MONOMER.
 CC -1- SUBCELLULAR LOCATION: Periplasmic.
 CC -1- SUBCELLULAR LOCATION: ACINETOBACTER CALCOACETICUS CONTAINS TWO DIFFERENT
 PQQ DEPENDENT ENZYMES WITH GDH ACTIVITY. GDH-A PREFERENCES
 2-DEOXYGLUCOSE AS SUBSTRATE, THE SPECIFIC SUBSTRATES FOR GDH-B ARE
 DISACCHARIDES.
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL PQQ DEHYDROGENASE FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation
 at the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: X07235; CAA30222.1; -.
 DR PIR: S00943; S00943.
 DR InterPro: IPR001479; Bac_PQQ.
 DR InterPro: IPR002372; Bac_PQQ_repeat.
 DR Pfam: PF01011; Bacterial_PQQ_7.
 DR PROSITE: PS00363; BACTERIAL_PQQ_1; 1.
 DR PROSITE: PS00364; BACTERIAL_PQQ_2; 1.
 DR Oxioreductase; PQQ; Transmembrane; Periplasmic; Signal.
 KW SIGNAL
 FT 1 34 801 POTENTIAL.
 FT CHAIN 1 34 801 GLUCOSE DEHYDROGENASE-A
 FT FT 34 801 (PIRLOQUINOLINE-QUINONE).
 FT TRANSMEM 39 55 POTENTIAL.
 FT TRANSMEM 59 79 POTENTIAL.
 FT TRANSMEM 94 108 POTENTIAL.
 FT TRANSMEM 119 138 POTENTIAL.
 FT ACT_SITE 471 471 BASE (POTENTIAL).
 FT SEQUENCE 801 AA; 86956 MW; 2F4EA222EB5450D8 CRC64;
 SQ

Query Match 12.0%; Score 323.5; DB 1; Length 801;
 Best Local Similarity 21.8%; Pred. No. 1.4e-16;
 Matches 152; Conservative 75; Mismatches 193; Indels 277; Gaps 31;

QY 20 PAFQAVTPTVDLLANPPGEMISYQONENRHSPLQITTEVNGOLQVWAGMGPK 79
 Db 152 PETAQAVPGVAE-----SDMPAYGRQAGVRSPLQINDQNVKLVKAWT--LATGD 202
 QY 80 V-----GVTPLEHDGVVYLANPDVIOADANTGDLWEHRQLPNIAITLNSF- 127
 Db 203 LKTDNDSGETTNOVTPKIGNNMFICTAHQOLAIIDPATGKEKW---FEDEKLKTDTSFQ 259
 QY 128 ----- 127
 Db 260 HLTGCVMYDANNTTEFATSLQSKSSSTQCPKRYEVPVNDGRIVANADTGKACDFG 319
 QY 128 -----GIWANG-VIVAGS--TCQYS---PFGCVSGHDSATG 158
 Db 320 QNGOVNIOEEMFAYPGGYMPTSPGIYTGSTVYIAGSVTDNSNKEPSG-VIRGYDVNTG 378
 QY 159 EELMRYEIPRAGE-----EGDETNGNDYEARMGTGANGQITYPVNLVHYGSTAVGP 212
 Db 379 KLLM--VEDTGAADPNAMPGSGTFFVNS-----DNAMAPLAYDKKLIV-YVPTGV-- 427
 QY 213 ASERTGRITPGGLTG-----TNTRAVVRPDGTGEIYVHRQLPRDNMDDECTFEEM 262

Db 428 -----GTP--DINGGDRTELKERVANSMLAINASGKLVNMFQTHDHLMDMDVPSQPS 479
 QY 263 VTNVQVQSTEMEGLOSINP--NAATGERVLTGP-----CKTG--TMOQFD 306
 Db 480 LADIKRAGQVPAIVLTKTGNAFVLDLR--NGOPTVYVTKRPVQYVKKRQPPQKGF 537
 QY 307 AETGEF-----LWA-----RDINYQNMIESIDENGIVTVNED 338
 Db 538 SKTQPSDLNAPQDKLTDKMDGATMLQDLMCRVSFKRLNIGIYTPSENGTL----- 592
 QY 339 AILKELDEVDCPTFLGGRDPSALNPS-----GIYFI----- 374
 Db 593 -----VFPNGLVFEMGMSVNPDRQYAVNPNLPGVPSRLIPADNRAQTFAG 641
 QY 375 -----PLNNQVDMADVQDEFTSMQVNTSNVTKLPKGMIGRIDAIDISTGRTLM 426
 Db 642 ACTEGVQPMYGPY-----GVEISAFSLPLGPKQOPAMGVAGVGLKTHEYVW 691
 QY 427 -----SVERAANYSPVLSGGVYLNQGT-DKTFALSGEGETLM 467
 Db 692 KKRIGTHDSLPLNLPQPAVKIGVPLGSGISITAGNVMPFGATQDNYLRAFNVTNGKILM 751
 QY 468 QTRLATVASGA--ISTEVDGMQYVAT-AGGCVSYGS 501
 Db 752 EARLP--AGGQATPMYTBINGQYVIMAGGHSFGT 786

RESULT 13
 DHG_GLUOX STANDARD; PRT; 808 AA.
 ID DHG_GLUOX
 AC P27175;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Glucose dehydrogenase [pyrroloquinoline-quinone] precursor
 DE (EC 1.1.99.17).
 GN Gdh.
 OS Gluconobacter oxydans (Gluconobacter suboxydans).
 OC Bacteria; Proteobacteria; alpha subdivision; Acetobacteraceae;
 OC Gluconobacter.
 OC NCBI_TaxID=442;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=92017653; PubMed=1833618;
 RA Cleton-Jansen A.-M., Dekker S., van de Putte P., Goosen N.;
 RT "A single amino acid substitution changes the substrate specificity
 of quinoprotein glucose dehydrogenase in Gluconobacter oxydans.";
 RL Mol. Gen. Genet. 229:206-212(1991).
 RN [2]
 RP REVISION TO 213.
 RA Goosen N.;
 RL Submitted (FEB-1992) to the EMBL/Genbank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: D-glucose + acceptor = D-glucono-1,5-lactone +
 reduced acceptor.
 CC -1- COFACTOR: PQQ.
 CC -1- SUBCELLULAR LOCATION: OUTSIDE OF THE PERIPLASMIC MEMBRANE.
 CC -1- MISCELLANEOUS: THE P1 FORM CAN OXIDIZE ONLY D-GLUCOSE, WHILE THE
 P2 FORM CAN ALSO OXIDIZE MALTOSE. THE SEQUENCE OF P1 FORM IS SHOWN
 HERE.
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL PQQ DEHYDROGENASE FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation
 at the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: X62710; CAA44594.1; ALT_SEQ.
 DR PIR: S17716; QPKEX.
 DR InterPro: IPR001479; Bac_PQQ.

DR InterPro: IPR002372; Bac_PQQ_repeat.
 DR Pfam: PF01011; Bacterial_PQQ_7.
 DR PROSITE: PS00363; BACTERIAL_PQQ_1; 1.
 DR PROSITE: PS00364; BACTERIAL_PQQ_2; 1.
 DR Oxidoreductase; PQQ; Transmembrane; Periplasmic; Signal.
 FT CHAIN 1 33 POTENTIAL.
 FT SIGNAL 34 808 POTENTIAL.
 FT TRANSMEM 35 54 GLUCOSE DEHYDROGENASE [PYRROLOQUINOLINE-
 FT TRANSMEM 59 76 QUINONE].
 FT TRANSMEM 94 108 POTENTIAL.
 FT TRANSMEM 123 138 POTENTIAL.
 FT ACT_SITE 470 470 BASE (POTENTIAL).
 FT VARIANT 788 788 H -> N (IN P2 FORM).
 SO SEQUENCE 808 AA; 87567 MW; 0F4160DA7652445 CRC64;

Query Match 10.9%; Score 295; DB 1; Length 808;
 Best Local Similarity 20.1%; Pred. No. 2e-14;
 Matches 141; Conservative 83; Mismatches 230; Indels 248; Gaps 25;

QY 13 ALALLAPAFAYTPVTEDELLA-----NPPAGMISYGGQENYRHSPLTQIT 61
 DB 128 AVLAFLFSLTPDHDISGELPQIYANASPADDPNVPASMAHAYGRTQAGDRWSPINQINA 187
 QY 62 ENYGOLOLV-----AKMOPGV--QVTPLIHDGVYLANPBDVYQALDAIDAKTGLIM 112
 DB 188 TNSNKLAVAHITKDMNSNDPGEQNTATPIEFNNMLYMSLHQKLEAVDAGANGK 247
 QY 113 EHRROD---PNIATLNSFGI-----YANGVIY----- 136
 DB 248 VYDPKQIINPGFQHLRCGVSFHETPRANAMSDGNAPAPLDDAKDISILPVDNGRLVEVDAD 307
 QY 137 AGSTC-----QYSPGCFV----- 150
 DB 308 TGTCTGCGFNNGEIDLRVNPQYTPPGQYEPSTPVYIDKLIANSATIDNGSVKQASGA 367
 QY 151 -SGHDSATGEELM---RNYFIPRAGEGDETWGNDYFARMGTAKMOIITDPYTNLVHY 205
 DB 368 TOAFDYVYTKRYVWDASNPNOJLPDSHVFPHNSNSIVS---STDANLNLVYI 422
 QY 206 GSTAVGPASFEYORGP---GGTLVGTNTRF---AVRPDGEIVYHQTLPYRDMDOEC 257
 DB 423 PNGV-----GTPDQMGDRDKDSERFAPGIALNALNDGKLAWCYQVHNDIMMEL 473
 QY 238 TFEAMVTNVDPSTEMEGIQSINPNATGERRVLTVPOCKTGMQFDETG-EPILMAR 316
 DB 474 PSQPSLYDYTKDKGLVPAIYA-----PRTGDIFVLDRTSKETIVPAP 517
 QY 317 DT-----NYOMTESIDENGIVTVNDAILELDV-----EYD 349
 DB 518 ETPVPGGAAPGDHRTSPQPMQSO---LTLRPKNPLNDSIDNGTIFIDQMECSIYFHLRIE 574
 QY 350 -----VCPFLGGRDMPASALNDPSGIYFPLNNVYDMMAYDQ----- 388
 DB 575 GPFPPLSKGLIFPGDLGFEWGLAVDQROVAFAANPLISLFPVSQVLPFRGPNLME 634
 QY 389 -----EFTSMYNTSVTKLIPCKDMI-----GRDAIDISGRFL 425
 DB 635 ENAKGTGEGTGLQHNHYGIPYAVNHLPLDVLPLFGIKMPCRTPRPGVYAGIDGLTKNVY 694
 QY 426 WSVERRAANVS-----PV-----LSTGGGVLE-NGCDTRYRALSOETG 463
 DB 695 WQHRNGTLRBSMTGSSILPLPIPIKIGVPSLGGPLSTAQNLFLTASMDYIYRAVNLMTG 754
 QY 464 ETLMQTRLATVASGOAISEVVDGMQYVAIAGGVSYGSLNS 505
 DB 755 KVLMDRLPAGAQATPTIYAINKQYI-----VYTAGGHS 790

RESULT 14
 QUITA_ACICA STANDARD: PRT; 809 AA.

AC Q59086;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Quinate/shikimate dehydrogenase [pyrroloquinoline-quinone]
 DE (EC 1.1.99.25) (MAD(P)-independent quinate dehydrogenase).
 GN QUITA.
 OS Acinetobacter calcoaceticus.
 OC Bacteria; Proteobacteria; gamma subdivision; Moraxellaceae;
 CC Acinetobacter.
 OX NCBI_Taxid=471;
 RN [1]
 RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
 RC STRAIN-BD413 / ADPL;
 RX MEDLINE=95095936; PubMed=8002591;
 RA Elsemore D.A., Ornstun L.N.;
 RT "The pca-pob supraprotein cluster of Acinetobacter calcoaceticus
 RT dehydrogenase";
 RT J. Bacteriol. 176:7659-7666(1994).
 RN [2]
 RP SEQUENCE OF 1-20 FROM N.A.
 RC STRAIN-BD413 / ADPL;
 RX MEDLINE=96011389; PubMed=7592351;
 RA Elsemore D.A., Ornstun L.N.;
 RT "Unusual ancestry of dehydrogenases associated with quinate catabolism
 RT in Acinetobacter calcoaceticus";
 RL J. Bacteriol. 177:5971-5978(1995).
 CC 1- FUNCTION: CAN ACT EITHER ON QUINATE OR ON SHIKIMATE.
 CC 1- CATALYTIC ACTIVITY: Quinate + pyrroloquinoline-quinone -> 5-
 CC dehydroquininate + reduced pyrroloquinoline-quinone.
 CC 1- CATALYTIC ACTIVITY: SHIKIMATE + PYRROLOQUINOLINE-QUINONE -> 3-
 CC DEHYDROSHIKIMATE + REDUCED PYRROLOQUINOLINE-QUINONE.
 CC 1- CORFACTOR: PQQ.
 CC 1- PATHWAY: QUINIC ACID CATABOLIC PATHWAY; FIRST STEP. THIS PATHWAY
 CC ALLOWS GROWTH OF BACTERIA WITH QUINATE BY ITS CONVERSION TO
 CC PROTOCATECHUATE AND SUBSEQUENT METABOLISM BY THE BETA-KETOADIPIRATE
 CC PATHWAY.
 CC 1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC 1- INDUCTION: BY PROTOCATECHUATE.
 CC 1- SIMILARITY: BELONGS TO THE BACTERIAL PQQ DEHYDROGENASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>
 CC or send an email to license@sdb.ch).
 CC
 DR EMBL: L05770; AAC37161.1;
 DR InterPro: IPR001479; Bac_PQQ.
 DR InterPro: IPR002372; Bac_PQQ_repeat.
 DR Pfam: PF01011; Bacterial_PQQ_7.
 DR PROSITE: PS00363; BACTERIAL_PQQ_1; FALSE_NEG.
 DR PROSITE: PS00364; BACTERIAL_PQQ_2; FALSE_NEG.
 DR Oxidoreductase; PQQ; Quinate metabolism; Transmembrane.
 FT TRANSMEM 14 34 POTENTIAL.
 FT TRANSMEM 41 61 POTENTIAL.
 FT TRANSMEM 68 88 POTENTIAL.
 FT TRANSMEM 90 110 POTENTIAL.
 FT TRANSMEM 127 147 POTENTIAL.
 SO SEQUENCE 809 AA; 88196 MW; 71F67CEBBA62BFCB CRC64;

Query Match 9.4%; Score 253; DB 1; Length 809;
 Best Local Similarity 21.5%; Pred. No. 2.8e-11;
 Matches 146; Conservative 68; Mismatches 245; Indels 220; Gaps 24;

QY 3 PSLIMASGALALLAA--PAFAQVTPVTEDELLANPAGMISYGGQENYRHSPLTQIT 61
 DB 148 PHEYVASGEELPLVVPFAKKV-----NMWHYGNDAAGGSFVALDDINR 193

Query Match	Best local Similarity	Score	DB 1	Length	790:
Matches	143: Conservative	20.38: Pident	No. 1,1e-09:	Mismatches	233: Indels
					259: Gaps
9	ASAGATALLA--APAAQVTPVDE---	LIANPPAG--EWISGONGENRRSPLOT	91	59	
110	ATVATLAAGAIGCMVPHPPVAGNAGMTAVPGSVQDQMSAIGTNDGGRFALDOI	169			
60	TTEVNGQLQVLMARQMGPKV---	QVPLIHDEVMVLANPDVIOAIDAKTGD	109		
170	NRSN-GRPAASGPGPTTPEGLANSDNGCAEDQITPLVQGEKVFCTPHNNLTALDASTGK	228			
110	LIMHRRQL---	PNATLNSFCIVANGVYVAGSGCQSP	145		
229	QIM--KREINATSSVMQRCRGLEYFDADALPAPVAMPSPIAV--	TYAQCANCRRL	283		
146	FGCFVSGH---DSATGEELMRYFLPRAGEGD---		175		
284	FNTIDGRLIIVADDTGA---	FCQGGSGNGVDLAKLGAAPDFYOLTSPLVAT	337		
176	---	ETW---GNQVE---ARMNTGAM	191		
338	TVVGGRTADDNVQTDMPGGVVRGSMWSPYRSAGIDGPNHDAQAPAAAGSSYVNSPTNW	397			
192	GOITTDVPTNLVHGSTRVAPASFTQGTGGTLYGTN---	TRPAVPRDGEI	241		
398	APMSYDAANNTVF---	LPFGGSTDLYGABERTALDHRYGASVALDATTGAE	446		
242	VNRHOTLPDMNDQCEFTFEMVNTVNDVQSTEMEGLOSTINPAAVAGERRYLTGPCKTGT	301			
447	KWYQYVHNDLMD---	FDLPQPSL--IDFPRQDSHTPAVYIGTAGQ	490		
302	MMQFDATGEELM-ARDT---	NY---QNMIESIDENGIVYVEDALIKEDYE	347		
491	IYVLDRATGRPLTEVREKVPYKGSIDIAHROYAPRPLSLVGMPOIGTKHILTESDMGATAMD	550			
348	YDVC---	PTFLGGRMPASALNPDGCIYF---	373		
551	QMLCRIFAQMRYEGLYTAPGTDSVLSPPGSLGGMNGGSLSTDPYHDVFAFADMRLGLNV	610			
374	--LPLAN---	YCYDMAVNDDEFISMVYNTSNVTL---	PRGDKMIGI	414	
611	QMIPADIRKKAAGGGBAYNTGKAVPLKGRPYAVNKKRFLSALGIPCOAPP---	YGTL	666		
415	DAIDISTGRTLSVERAAN---	YSYV---	LSTGGVLYFGNGT-DRY	454	
667	SAIDKTSIMQVPGVTQDTGPFQIKMHLPIPIGMFTLGGLTSTGGVLYFIAAGTQDY	726			

walick-934-128.pwp.rsp

Search completed: May 24, 2002, 10:30:54
Job time: 632 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2002, 10:30:17 ; Search time 93.8 Seconds
(without alignments)
931.369 Million cell updates/sec

Title: WALICK-934-128.PEP
Perfect score: 2705
Sequence: 1 MKPSLWASAGALALIAAP.....GMQYVAAGGVSYSGSLNS 505

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: SPTRMBL_19:*
2: sp.archaea:*
3: sp.bacteria:*
4: sp.fungi:*
5: sp.human:*
6: sp.invertebrate:*
7: sp.mhc:*
8: sp.mammal:*
9: sp.mmc:*
10: sp.organelle:*
11: sp.phage:*
12: sp.plant:*
13: sp.podent:*
14: sp.virus:*
15: sp.verticillate:*
16: sp.unclassified:*
17: sp.rv1:*
18: sp.rv2:*
19: sp.rv3:*
20: sp.rv4:*
21: sp.rv5:*
22: sp.rv6:*
23: sp.rv7:*
24: sp.rv8:*
25: sp.rv9:*
26: sp.rv10:*
27: sp.rv11:*
28: sp.rv12:*
29: sp.rv13:*
30: sp.rv14:*
31: sp.rv15:*
32: sp.rv16:*
33: sp.rv17:*
34: sp.rv18:*
35: sp.rv19:*
36: sp.rv20:*
37: sp.rv21:*
38: sp.rv22:*
39: sp.rv23:*
40: sp.rv24:*
41: sp.rv25:*
42: sp.rv26:*
43: sp.rv27:*
44: sp.rv28:*
45: sp.rv29:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1500.5	55.5	608	2	093RE9
2	523	19.3	698	2	09KH03
3	489.5	18.1	742	2	053362
4	481	17.8	708	2	046444
5	478.5	17.7	601	2	09EYH8
6	477.5	17.7	601	16	09EYH9
7	474	17.5	695	2	09EYH9
8	470	17.4	629	2	09AQ08
9	461	17.0	601	2	09EYH9
10	459	17.0	601	2	09EYH9
11	430	15.9	633	2	09EYH9
12	422.5	15.6	623	2	09EYH9
13	417.5	15.4	623	2	09EYH9
14	408	15.1	695	2	09EYH9
15	396	14.6	573	2	09EYH9
16	302	11.2	790	2	09X255

17	282.5	10.4	785	16	098KF6
18	277	10.2	786	2	P95466
19	273.5	10.1	803	16	091115
20	258.5	9.6	777	16	092RB3
21	232	8.6	644	2	052551
22	227.5	8.4	639	2	P77931
23	208.5	7.7	470	2	032699
24	206	7.6	470	2	032699
25	203.5	7.5	180	2	032697
26	200.5	7.4	180	2	032696
27	200.5	7.4	180	2	032700
28	200.5	7.4	181	2	032615
29	199.5	7.3	181	2	032692
30	197.5	7.3	181	2	032703
31	196.5	7.3	181	2	032621
32	195.5	7.2	182	2	032706
33	193.5	7.2	180	2	032701
34	191.5	7.1	182	2	032612
35	191.5	7.1	182	2	032693
36	190.5	7.0	172	2	033894
37	186.5	6.9	172	2	033882
38	186.5	6.9	179	2	032707
39	185.5	6.9	180	2	032694
40	185.5	6.9	180	2	032704
41	185	6.8	171	2	033884
42	183.5	6.8	179	2	032702
43	182.5	6.7	184	2	032695
44	181.5	6.7	185	2	0918K8
45	176.5	6.5	172	2	093K58

ALIGNMENTS

RESULT 1
ID 093RE9 PRELIMINARY; PRT; 608 AA.
AC 093RE9;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE ALCOHOL DEHYDROGENASE.
GN ADH.
OS Pseudoglucobacter saccharotogenes.
OC Bacteria; Pseudoglucobacter.
OX NCBI_TaxID=133921;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IFO 14464;
RC Shiba T., Saito Y.,
RT "Alcohol dehydrogenase."
RT Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB046580; BAB62258.1; -
DR SEQUENCE 608 AA; 65101 MW; 0ACCE97AE11BA570 CRC64;

Query Match	55.5%	Score 1500.5;	DB 2;	Length 608;
Best Local Similarity	50.3%	Pred. No. 2.8e-101;		
Matches 293;	Conservative 63;	Mismatches 139;	Indels 87;	Gaps 6;
DB	10	SAGALALIAAPAFQ-----	VTPTDELNPAGWEISYGOQEN	50
DB	16	STALIASLSPAPADHANAAPKAGASAIENFQPTADLACKNANPILRGVNG		75
DB	51	YRHSPLQITTEVNGGOLVWARGKQVYPLIHDSVYLANPGVYQAIADKQGL		110
DB	76	WGSPPLQINKVGDGLQVWYRTEPGSEGAIAVNGVIFLGNITNDVIAIDKQSL		135
DB	111	IMEHROLPNAT-LNSFG-----		128
DB	136	IMEYRRLPSASKINSIGAKRSIALFGDKYFVSNDVFNALDAKTKLAWETNRGQ		195
DB	129	IVANGTVAGSFCQYSPFCFVSGHDSATGELMKNFTIPRAGEGDET		177

```

Db 196 VEEGVANSSGPIVDVGVAGVSTGCGECYVTGTDAESGELMKNFIPRPGEGDIT 255
Oy 178 MGN-DYKRWKMGWQIYDPTVNLVHGSFVAVGPAFSGTGTGGLXGNTREAVRP 236
Db 256 WGGAPYNNRWMTGAGWQIYDPELDLYVYGSFAGAPASVQAGTGGTLAGINTREAVRP 315
Oy 237 DTGEIWRHQTLPDRMDQCEPEMAYTVNDVQPSFTEMBGLQSLNPNAATGE-RRVLGV 295
Db 316 KTGEVVRKHQTLPDRMDQCEPEMAYTVNDVQPSFTEMBGLQSLNPNAATGE-RRVLGV 375
Oy 296 PCKGTGMQFPAETGEFLMARDTNTQNMIESIDENGIYTVNEDAILKELDEVYCTEL 355
Db 376 PCKGTGMQFPAETGEFLMARDTNTQNMIESIDENGIYTVNEDAILKELDEVYCTEL 435
Oy 356 GGRDMPGSAALNDPSGIFPLNNVCYDMAVDOFTSDMYNTSNVTKLPKGMIGRID 415
Db 436 GGRDMPGSAALNDPSGIFPLNNVCYDMAVDOFTSDMYNTSNVTKLPKGMIGRID 495
Oy 416 AIDISGRTLMSEVERAANYSFVLTGGVLENGTDRYFRALSOETGETLMOTRLATYA 475
Db 496 AIDISGRTLMSEVERAANYSFVLTGGVLENGTDRYFRALSOETGETLMOTRLATYA 555
Oy 476 SCQATSYEVDSQVYVAGG-----VYSGSLN 504
Db 556 SCQATSYEVDSQVYVAGG-----VYSGSLN 597

```

RESULT 2

```

O9KH03 PRELIMINARY; PRT; 698 AA.
AC O9KH03;

```

```

Dt 01-OCT-2000 (Tremblrel. 15, Created)
Dt 01-OCT-2000 (Tremblrel. 15, Last sequence update)
Dt 01-DEC-2001 (Tremblrel. 19, Last annotation update)
De TERAHYDROFURFURYL ALCOHOL DEHYDROGENASE.
Os Alcaligenes eutrophus (Ralstonia eutropha).
Oc Bacteria; Proteobacteria; beta subdivision; Ralstonia group.
Ox NCBI_TaxID=510;
Rn [1]
Rp SEQUENCE FROM N.A.
Ra MEDLINE=21125557; PubMed=11222593;
Rt Zarnit G., Schrader T., Andreesen J.R.;
Rt "Catalytic and Molecular Properties of the Quinohemoprotein
Rt Tetrahydrofurfuryl Alcohol Dehydrogenase from Ralstonia eutropha
Rt Strain B0."
Rt J. Bacteriol. 183:1954-1960(2001).
Dr EMBL: AF277373; AAF86335.1;
Dr HSP: Q9Z477; IFLG.
Dr InterPro: IPR002372; Bac_PQO_repeat.
Dr InterPro: IPR000345; CytC_heme_Bind.
Dr Pfam: PF003088; CytC_C1.
Dr Pfam: PF01011; Bacterial_PQO_6.
Dr PROSITE: PS00364; CYTOCHROME_C_1.
Dr PROSITE: PS00190; CYTOCHROME_C; UNKNOWN_1.
Dr SEQUENCE 698 AA; 75471 MW; 9C87F180DB1D2FB2 CRC64;

```

Query Match Best Local Similarity 19.3%; Score 523; DB 2; Length 698; Matches 161; Conservative 82; Mismatches 216; Indels 126; Gaps 22;

```

Oy 10 SAGALALAAPAF--QVPTVDLLANPPAG--EWISGQNOENYRHSPLQTITTEVNG 65
Db 14 AAASVALPAPMARGANAAARVDAIRANEASTPMPSPYGDVYETRFSLKLEQVNAQNR 73
Oy 66 QIDLVARQGPCK-VQVPLIHGVMYLANPGVYQAIIDAKTGDLLMEHRQLPN----- 120
Db 74 NLGLAMSYDLESTRGEATPLVVDGVMTYSAFWSVHAIDARTKRLMTYDPOVPRQDAY 133
Oy 121 -----IA-----TNSFGIYA 131

```

```

Db 134 KCCQDVNVNKGVALYQGVFVGAFGDRLVAIDATGKKWEDDTITDRSKSYTITGAPRY 193
Oy 132 NG-VIVASTOYSPFGCFVSGHESATGELMKNYFIP-----RAGEGDEFTMGND 181
Db 194 NCKVLIIGNGAGLETGVRG-VITAYDAETGKQOQNMWYVTPGDPARPFENEMAAATW--D 250
Oy 192 YEAR-WMTG-----AMQIYDPTVNLVHGSFVAVGPAFSGTGTGGLXGNTREAVR 235
Db 251 PSKGWINGGGGGVWNTMARDPELNLMTIGTGAGWSKRLSPKGGNLIYAASV-VALN_309
Oy 236 PDTEIWRHQTLPDRMDQCEPEMAYTVNDVQPSFTEMBGLQSLNPNAATGE-RRVL 293
Db 310 PDTEIWRHQTLPDRMDQCEPEMAYTVNDVQPSFTEMBGLQSLNPNAATGE-RRVL 351
Oy 294 GVPCKTGMQFPAETGEFLMARDTNTQNMIESIDENGIYTVNEDAILKELDEVYCTEL 351
Db 352 HAP-KNGFFVDIRTNGKFLSAKNFVDVWMAAGYCKNRPVETPOADTSGKPA----DVV 406
Oy 352 PTFEGGRDMPGSAALNDPSGIFPLNNVCYDMAVDOFTSDMYNTSNVTKLPKGMIGRID 400
Db 407 GGPFGAHNMHSMSEFKLGLAFIPQAHVPL-TLADNKEW-----VHNOKDSPEAHRGVGN 461
Oy 401 -----VTKLPKGMIGRIDAIDISTGRTLMSEVERAANYSFVLTGGVLENGTDRYFR 456
Db 462 LGMLVNAEPPRSKPMGRILAMPDLAOKAVMHRDAGPWNNGITLATGNTLVFGTADGRLV 521
Oy 457 ALSOETGETLMOTRLATYASGOATSYEVDSQVYVAGG-----VYSGSLN 500
Db 522 AYHAATGKELMQAPRGSGVVAAPVYTLDDGROYSAVAVGNGVYG 566

```

RESULT 3

```

O53362 PRELIMINARY; PRT; 742 AA.
AC O53362; Q44159;

```

```

Dt 01-NOV-1996 (Tremblrel. 01, Created)
Dt 01-NOV-1996 (Tremblrel. 01, Last sequence update)
Dt 01-OCT-2001 (Tremblrel. 18, Last annotation update)
De DEHYDROGENASE SUBUNIT OF ALCOHOL DEHYDROGENASE.
Os Acetobacter pasteurianus (Acetobacter turbidans).
Oc Bacteria; Proteobacteria; alpha subdivision; Acetobacteraceae;
Ox NCBI_TaxID=438;
Rn [1]
Rp SEQUENCE FROM N.A.
Ra STRAIN=NC11380;
Rt MEDLINE=94042848; PubMed=8226628;
Rt Takemura H., Kondo K., Horiouchi S., Beppu T.;
Rt "Induction by ethanol of alcohol dehydrogenase activity in Acetobacter
Rt pasteurianus."
Rt J. Bacteriol. 175:6857-6866(1993).
Dr EMBL: D13893; BAA40252.1;
Dr HSP: Q9Z477; IFLG.
Dr InterPro: IPR001479; Bac_PQO.
Dr InterPro: IPR002372; Bac_PQO_repeat.
Dr Pfam: PF01011; Bacterial_PQO_6.
Dr PROSITE: PS00363; BACTERIAL_PQO_1;
Dr PROSITE: PS00364; BACTERIAL_PQO_2;
Dr SEQUENCE 742 AA; 81680 MW; 031C4E9B6B13EC98 CRC64;

```

Query Match Best Local Similarity 18.1%; Score 489.5; DB 2; Length 742; Matches 162; Conservative 83; Mismatches 215; Indels 147; Gaps 22;

```

Oy 11 AGALALAAPAFQVPTVDLLANPPAG--EWISGQNOENYRHSPLQTITTEVNGQIQ 68
Db 23 AAALPYAAVAPARADGCGTGEALIHADHPKEMLSYGRYSQYSPIDQINRSNVGDLK 82
Oy 69 LVMARQGPCKVQ-VQVPLIHGVMYLANPGVYQAIIDAKTGDLLMEHRQLP-NIA----- 122
Db 83 LAMVYTLDTNRGQEATPLVVDGVMTYSAFWSVHAIDARTKRLMTYDPOVPRQDAY 142

```

123 ---TLN-----SFGI----- 129
 143 CDTVNNAGAGYNNKVFNGTFFDGLVLAADATGKWEVNTIPADSLGKORRYTVDGAVR 202
 130 VANGVAVAGSTCOYSPFGC--FVSGHOSATGEELMRYEIPRAGEBD----- 175
 203 VAKGLVILN--GSGEPGARGFSAFPADEICKLMKRYEYTPNNKNEPDHADVNLMSKA 260
 176 -ETMGNDYEARMWT-----GANGQITDPVTLNLYHGSTAVGASSETORCPGTLY-- 226
 261 YKTMGP--KAMVWROGGGTVWDSLVDPVSDLY--LAAG-----NSPNNYKASE 309
 227 --GTN-----TRFVRPDTEIYWRHQLPRDNWDECTFEMMTVNDVOPSTEMEGLQSI 280
 310 GIGSNLGLSIVALKPEPTGEYVHHQAPPMQOWDYSVOQLMTLDMV----- 357
 281 NPNMATGE--RRVLGVPCKTGMQPDATGEFLMARDPNTQNMIESIDE--NGIYVNEED 338
 338 -----NGEMRIVIHAP--KNGFFYVLDARTGEFLAGKNIVYONNANGIDPLTGRPIYND 411
 339 ALLKELDEVYCTFLGGRDMPSAALNPDGIFYPLNNVCY-----DMMA 385
 412 GYTLTGKFWGIGEPPLGSAHNMGMASPKTHLVYLPALQIFGVYKNOVGGRPHPDAMN 471
 386 VDOEFTSMQVYTSNWTLPKPKMIRIDALISTGRITLMSVERAANYSPLSTGGGV 445
 472 VGLDMTKNGLPDPTE--ARTAYIKDLHGMLLAWDPKMEIYWKLDHGPWNGVLATGGDL 530
 446 LFPNGDRFRFALSOETGLMOTRLATVASGAISTEVDGMQYV-----IAGG 495
 531 LFGGLANGEFHAYDANNSDLYKFDAGSGLIASPYSVNGKOYVAVEVWGSIYPISSNG 590
 496 GVSYSGS 502
 591 GVRTSG 597
 RESULT 4
 Q46444 PRELIMINARY: PRT: 708 AA.
 ID Q46444
 AC Q46444
 DT 01-NOV-1998 (TREMBLER, 08, Last sequence update)
 DT 01-NOV-1998 (TREMBLER, 08, Last sequence update)
 DT 01-DEC-2001 (TREMBLER, 19, Last annotation update)
 DE QUTNOHAEMOPROTEIN ETHANOL DEHYDROGENASE TYPE I PRECURSOR
 (EC 1.1.99.-) (OH-BDH1).
 GN QHEBH.
 OS Comamonas testosteroni (Pseudomonas testosteroni).
 OC Bacteria; Proteobacteria; beta subdivision; Comamonadaceae; Comamonas.
 OX NCBI_TaxID=285;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 15667;
 RX MEDLINE=96184549; PubMed=8654419;
 RA Stoovogel J., Kraayveld D.E., van Sluis C.A., Jongejan J.A.,
 De Vries S., Duine J.A.;
 RT "Characterization of the gene encoding quinoxaloproduct ethanol
 dehydrogenase of Comamonas testosteroni";
 RT Eur. J. Biochem. 235:690-698(1996).
 RL [2]
 RN SEQUENCE OF 32-54 AND 477-490.
 RP STRAIN=ATCC 15667;
 RX MEDLINE=95324580; PubMed=7601151;
 RA De Jong G.A.H., Geerlof A., Stoovogel J., Jongejan J.A., De Vries S.,
 Duine J.A.;
 RT "Quinoxaloproduct ethanol dehydrogenase from Comamonas testosteroni.
 RT purification, characterization, and reconstitution of the apoenzyme
 with pyridoxal phosphate analogues";
 RL Eur. J. Biochem. 230:899-905(1996).
 RN [3]
 RP CHARACTERIZATION.
 RX MEDLINE=86242113; PubMed=3521592;
 RA Groen B.W., van Kleef M.A., Duine J.A.;

RT "Quinoxaloproduct ethanol dehydrogenase apoenzyme from Pseudomonas
 RT testosteroni";
 RL Biochem. J. 234:611-615(1986).
 RN [4]
 RP MEDLINE=21536088; PubMed=11679760;
 RX Oubrie A., Huijzinga E.G., Rozeboom H.J., Kalk K.H., de Jong G.A.H.,
 Duine J.A., Dijkstra B.W.;
 RA "Crystallization of quinoxaloproduct ethanol dehydrogenase from
 RT Comamonas testosteroni: crystals with unique optical properties";
 RL Acta Crystallogr. D 57:1732-1734(2001).
 CC -1- FUNCTION: CATALYZES THE DYE-LINKED OXIDATION OF PRIMARY ALCOHOLS
 CC TO THE CORRESPONDING ALDEHYDES AND THE (SUBSEQUENT) OXIDATION OF
 CC THE ALDEHYDES TO CARBOXYLIC ACIDS.
 CC -1- COFACTOR: POO, HEME, AND CALCIUM.
 CC -1- SUBUNIT: MONOMER.
 CC -1- SUBCELLULAR LOCATION: PERIPLASMIC (POTENTIAL).
 CC -1- INDUCTION: BY ETHANOL AND BUTANOL.
 CC -1- SIMILARITY: TO ALCOHOL DEHYDROGENASE (ACCEPTOR) PROTEINS.
 DR EMBL: X81880; CAA57464.1; -.
 DR HSP: 0924J7; 1FLG.
 DR InterPro: IPR002372; Bac_POO_repeat.
 DR InterPro: IPR003088; Cyt_C1.
 DR InterPro: IPR003329; Cyt_C1C.
 DR Pfam: PF01011; Bacterial_POO; 6.
 DR Pfam: PF00034; Cytochrome_c; 1.
 DR PRINTS: PR00605; CYTOCHROME_C1.
 DR Signal: POO; Heme; Calcium; Oxidoreductase; Periplasmic.
 KW SIGNAL: 1 31
 FT CHAIN: 32 708
 FT TYPE 1.
 FT BINDING: 635 635 HEME (COVALENT) (BY SIMILARITY).
 FT BINDING: 638 638 HEME (COVALENT) (BY SIMILARITY).
 FT METAL: 639 639 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
 FT METAL: 639 639 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
 SQ SEQUENCE 708 AA; 76822 MW; 99AB54BDD6ACAB3 CRC64;

Query Match 17.8%; Score 481; DB 2; Length 708;
 Best Local Similarity 25.4%; Pred. No. 9, 6e-27;
 Matches 150; Conservative 84; Mismatches 239; Indels 118; Gaps 19;
 3 PTLNLSASAGATALLAFAFAVTPVT-----DELANPP-AGEMISYGNQEN 50
 12 PGRVWVLLAACLG--SNAFACTGPAQAQAAAVORVDDFLRAVAARTPMTITGVDAE 69
 51 YRHSPLQITTEVNGQLVWARGMPGK-YQVPLHDGVVLANPGVYIQAIDAKTGD 109
 70 TRYSRLDQINAAVKNKIDGLAMSYNLESTGVEATPVVDGIMVYASWSVHAIDRTGN 129
 110 LWEHRRLDPN-----IA----- 122
 130 RIWYDPOIDSTGFKCCDVNRGVALMKKYYVAGMDGLIALDAATGKWHOMTFE 169
 123 -----TLNSFGIYANGVYAGST-CQYSPFGCVSGHDSATGEELMRYEIP----- 168
 190 GKGSLITGAPRFKGVIIIGKRAEYVGR-YITAYAEIGERKWRFSVGPSPKRF 248
 169 --RAGEGDEFTWNGDYEARMTGA-----WGQITDPVTLNLYHGSTAVGASSETORCPGT 221
 249 EDESMKRAATM--DPSCKWEDAGGGGTMDSMFTDELMTMTVVGKNSPMSHVRSKP 306
 222 GGLYGTNTRFAVRPDTEIYWRHQLPRDNWDECTFEMMTVNDVOPSTEMEGLQSI 281
 307 GGDNLVLAIVLADPPTGKYKMHYETPGDNMDYTSQPMILADIKI----- 353
 282 PNATGRRVLTVPCKTGMQPDATGEFLMARDPNTQNMIESIDE--IYVNEEDAI 340
 354 ---AGRKAKYILHAP--KNGFFYVLDRTNKGFLSAKNVYPVWASGDKKGPPIGI---AA 406
 341 LKELDEVYCTFLGGRDMPSAALNPDGIFYPLNNVCYDMAVQD-EFT----- 391
 407 ARDSKPDVAVPGYGAHNNHPSFNPQRTGLYLVLAQNVVAVNLDDKKWEPNAGRCKPQ 466

QY 392 SMDVNTSNVTKL-PPCKMIGRIDALDSTGTMSVERAANYSPLSTGGVLFNG 450
 Db 467 SGTGMNNAKFPENRPPSKRPRLALMDPVAAKAMSEVHSPMNGITLTAGVNVFQST 526
 QY 451 TDRYFRLSDEGTETLMOTRLATVASGQAISEYVDQMAYALA-GGVSVSG 500
 Db 527 ADGLVAVYHAHATGEKLEAPGTGVAAAPSTYMDGRQYVAVANGMGVIG 577

RESULT 5
 Q92YW8 PRELIMINARY: PRT: 601 AA.
 ID Q92YW8
 AC Q92YW8
 DT 01-MAR-2001 (Tremblrel. 16, Created)
 DT 01-DEC-2001 (Tremblrel. 16, Last sequence update)
 DE 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 GN KMAP
 OS Rhizobium meliloti (Sinorhizobium meliloti).
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Rhizobiaceae; Sinorhizobium.
 OX NCBI_TaxID=382;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=RM1021;
 RA Fennel B.J., Tiwari R.P., Dilworth M.J.
 RT "Regulation of C1 assimilation in Sinorhizobium meliloti."
 RL Submitted (SEP-2000) to the EMBL/Genbank/DBJ databases.
 DR HEMP; AF039488; AAC31643.1;
 DR HEMP; P38539; AAAH.
 DR InterPro: IPR002372: Bac_PQQ_repeat.
 DR Pfam: PF01011: Bacterial_PQQ; 7.
 SQ SEQUENCE 601 AA; 65866 MW; 09F743726B2E2A83 CRC64;

Query Match 17.7%; Score 478.5; DB 2; Length 601;
 Best Local Similarity 25.8%; Pred. No. 1.2e-26;
 Matches 155; Conservative 90; Mismatches 210; Indels 145; Gaps 24;

QY 14 LALLAFAFAQVTPVDEL-LANPAGEMISGQNGENRHSPLTOTTENVGOLQVW 71
 Db 8 LAIMISGGAQVAFANDELQKLIIDP-NQMAIQTDYANLRSKLDQINKDKVGLQVW 66
 QY 72 --ARGMOPGKQVQVPLIHGQVMTLANP-GDVIAQID-AKTGDLMEHR-ROLPNITLNS 126
 Db 67 TFSIGVLRGH-EGSPLYIGDLMTVHTPPNTVYALDLSKQGLVWKYEPKODPNVIVMC 125
 QY 127 FGIYANGVIYAG-----STCOYSP----- 145
 Db 126 CDIVNRCVAAADNKIFLHQADTTVALDAKTGKVIYSKNGDATKGETNTATVMPYDKI 185
 QY 146 -----FGC--FVSGHDSATGEELMRNYFT-----PRAGEGDET 177
 Db 186 LVGISGGEFVGRGHVATYASMDGKVLKMGYSKGPDSOTLIDPEKTHLQKPVKDSGLT 245
 QY 178 WGNDEYARMMTG---AMGOITDPVTNLVYHGSTAVGPASETOGTGGTLYTNTFRV 234
 Db 246 WEGD---QMKIGGCTTWGYSYDPEENLVYGTGNPSTWPTOR--PGDNRM-SMTIFAR 299
 QY 235 RPDTEIYWRHQTLPRLNMDQECTFEEMVNTVNDVOPSTEMEGLOSINPNAATGERRVLT 294
 Db 300 DVDITGMAKMLYQMTPHDEMVDYDGVNEMILTGQIDGK-----DRKLLTH 343
 QY 295 VPCKTGIMQFPAETGEFL-----WADRTNYQNMIE-----SIDENGIV 333
 Db 344 FD-RNGFGYTMRYTDELVAEKYDPTVMNATEVMDPDKSGRQVAVQYSTEDNG-- 400
 QY 334 TVNEAAILKELDEYDVCPFLGGRDMPSAALNPDGSIYFIPLNNVCYDMAVAQDEFTSM 393
 Db 401 ---EDNTNT-----GVCPAALGTCKDQOPAAVSPKTELFYVPTNHCMDYEPFVSYTAG 451
 QY 394 DVYNTSNVTKLPQKDM---IGRIDALDSTGTMSVERAANYSPLSTGGVLFNG 450

Db 452 QPYVGAITLSKYP-KDSHGCMGNFTACDNKESGKIKMSLPEPFSWSGALATAGDVYFCT 510
 QY 451 TDRYFRLSDEGTETLMOTRLATVASGQAISEYVDQMAYALG-----GVSYSGLNS 505
 Db 511 LESTYLAADVAADGKELVREFKTPSGVIGVMTYANREGQYVAVLVGGVAGIGLAIGLITN 570

RESULT 6
 Q92WY9 PRELIMINARY: PRT: 601 AA.
 ID Q92WY9
 AC Q92WY9
 DT 01-DEC-2001 (Tremblrel. 19, Created)
 DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
 DE 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 GN PUTATIVE METHANOL DEHYDROGENASE PROTEIN, LARGE SUBUNIT (EC 1.1.99.8).
 GN SMD20173.
 OS Rhizobium meliloti (Sinorhizobium meliloti).
 OG Plasmid PSYMB (megaplasmid 2).
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Rhizobiaceae; Sinorhizobium.
 OX NCBI_TaxID=382;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=1021;
 RX MEDLINE=21396508; PubMed=11481431;
 RA Flann T.M., Weidner S., Wong K., Buhmester J., Chain P.,
 RA Vorheeler F.J., Hernandez-Lucas I., Becker A., Cowie A., Guzy J.,
 RA Golding B., Piehler A.,
 RT "The complete sequence of the 1,683-kb psymb megaplasmid from the N2-
 RL Proc. Natl. Acad. Sci. U.S.A. 98:9889-9894(2001).
 DR EMBL; AL603642; CAC48573.1;
 KW Oxidoreductase; Plasmid; Hypothetical protein; Complete proteome.
 SQ SEQUENCE 601 AA; 65759 MW; D73424FEED13ADB8 CRC64;

Query Match 17.7%; Score 477.5; DB 16; Length 601;
 Best Local Similarity 25.8%; Pred. No. 1.4e-26;
 Matches 155; Conservative 90; Mismatches 210; Indels 145; Gaps 24;

QY 14 LALLAFAFAQVTPVDEL-LANPAGEMISGQNGENRHSPLTOTTENVGOLQVW 71
 Db 8 LAIMISGGAQVAFANDELQKLIIDP-NQMAIQTDYANLRSKLDQINKDKVGLQVW 66
 QY 72 --ARGMOPGKQVQVPLIHGQVMTLANP-GDVIAQID-AKTGDLMEHR-ROLPNITLNS 126
 Db 67 TFSIGVLRGH-EGSPLYIGDLMTVHTPPNTVYALDLSKQGLVWKYEPKODPNVIVMC 125
 QY 127 FGIYANGVIYAG-----STCOYSP----- 145
 Db 126 CDIVNRCVAAADNKIFLHQADTTVALDAKTGKVIYSKNGDATKGETNTATVMPYDKI 185
 QY 146 -----FGC--FVSGHDSATGEELMRNYFT-----PRAGEGDET 177
 Db 186 LVGISGGEFVGRGHVATYASMDGKVLKMGYSKGPDSOTLIDPEKTHLQKPVKDSGLT 245
 QY 178 WGNDEYARMMTG---AMGOITDPVTNLVYHGSTAVGPASETOGTGGTLYTNTFRV 234
 Db 246 WEGD---QMKIGGCTTWGYSYDPEENLVYGTGNPSTWPTOR--PGDNRM-SMTIFAR 299
 QY 235 RPDTEIYWRHQTLPRLNMDQECTFEEMVNTVNDVOPSTEMEGLOSINPNAATGERRVLT 294
 Db 300 DVDITGMAKMLYQMTPHDEMVDYDGVNEMILTEQIDGK-----DRKLLTH 343
 QY 295 VPCKTGIMQFPAETGEFL-----WADRTNYQNMIE-----SIDENGIV 333
 Db 344 FD-RNGFGYTMRYTDELVAEKYDPTVMNATEVMDPDKSGRQVAVQYSTEDNG-- 400
 QY 334 TVNEAAILKELDEYDVCPFLGGRDMPSAALNPDGSIYFIPLNNVCYDMAVAQDEFTSM 393
 Db 401 ---EDNTNT-----GVCPAALGTCKDQOPAAVSPKTELFYVPTNHCMDYEPFVSYTAG 451

```

Fri May 24 1968
OY      394  DYVNTSNVTKLPCKDM--IGRIDIDISTGRTLSVERAANTSPVLTSTGGVLNCG 450
      OY      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      Db      452  QRYVATLISMP-KSHCGKGNFIANDRKRKIRMSLPEPFVSWGALATVAGDVEYGT 510
      OY      451  TDYRPAALSOETGETLMOTRLATVASGALISYEVNDGMYVIAAG-----GVSYGSLNS 505
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      Db      451  TDYRPAALSOETGETLMOTRLATVASGALISYEVNDGMYVIAAG-----GVSYGSLNS 505
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      Db      511  LQGYLAKVADPAKGLTLYRFTYPSGVIGNMTYAREKQIVAVILSGVGWAGILAGLTLN
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

[illegible]

```

ep.ispt
361 pSaALNDSCiITiPLiNVCYDMaNAVDEFTSMDYNT---SNVTKLPgKdMI-GRD 415
QY      | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
409 HSMsENPOTGLMTiPYOEiLPGVYRNREGsFKRKiDGLNtGTSDFtHEiP--RDVSGAIL 466
Db      | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
QY      | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
416 AIdISTGRtLMSYEKAANtSPYLtSGGVLtFNGSDtRFAiLsOEtGELtMOTRLAYa 472
      | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
467 AMDPVROREAMNPHSPFWNGtLStAGNLtVPGtAGOLHAVSADKORLMSPAOtGI 522
Db      | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
QY      | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
476 SGOLtSYEVNDsKOYVAiAG 495
      | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
527 VAAPiSFSLDGQIVAMAG 546
Db      | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |

```

[illegible]

Fri May 24 11:27:31 2002

walick-934-128.pcp.ispt

Query Match
Best Local Similarity 17.0%; Score 461; DB 2; Length 601;
Matches 153; Conservative 102; Mismatches 218; Indels 124; Gaps 26;

Db 1 MKPTSLMASAGALLAARAFAYOVPYIDELLANPAGEMISTGONENRHSPLTOT 60
1 MRVHLLALGAGLAA-ASBALANESYLKG--VAAP--ABOVLQTVYANTRIKSLIDIN 54
61 TENVOGLQVW--ARGMOPKGYOVTPLIHGVMYLANP-GDVOIADAKTG-DLIMHR- 115
55 ASVAKMLQVAVTFSTGVLKCH-EGSPLVGVNIMVHTPPFNITVYALDLDQAKIVKPEP 113
116 ROLPNA-----TLNFTGIVANGYIV-----AGS 139
114 KODPSVLPVWCCTVNRGLAYADGAILLNOADTVLSIDANSGVWVSKNDPSKGETN 173
140 TCQYSP-----FC--CPVSGSDATSGEELMRVYF----- 167
174 TATVLPVADKYIVGISGEGEVQCHVTATIDKSGKVMWNGISIGPDOLLIVDEKTTSLG 233
168 -PRAEEDDETWCNDYEAAMMG--AMGITYDDPTNLVHYGSTAVGASSTOQGTGCG 223
234 KPIAKDSLKTWEGD--OMKGTGGCTKMGWSTYDPKIDLMKATGSG--NPSIWNKORPFG 288
224 TLGTITREAVAPPIGELIVNRHDTLPBGNDDECTFPMVNTVNDVOSTEMEGLOSINPN 283
289 NKW-SMTIWMANBDTOMAKWYOMTPHDEWDFDGINEMILTD-----KRFDC----- 334

RESULT 9
P71509 PRELIMINARY; PRT; 601 AA.
AC P71509
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
GN MAFANOL DEHYDROGENASE LARGE SUBUNIT HOMOLOG.
OS Methylbacterium extorquens
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
NCBI_TaxID=408; Methylbacterium.
RP SEQUENCE FROM N.A.
RX MEDIAN-AM1;
RA Christensen-97312011; PubMed-9168622;
RT "Moleseridova L.; Lidstrom M.E.;
KT Molecular and mutational analysis of a DNA region separating two
RL EMBL; 072662; AAB8890.1; clusters in Methylbacterium extorquens AM1";
DR HSP; 038539; 44848
DR InterPro; IPR002372; Bac_P00_repeat.
SQ SEQUENCE 601 AA; 64952 MW; 68E4SC7059CB239 CRC64;

Query Match
Best Local Similarity 17.0%; Score 459; DB 2; Length 691;
Matches 145; Conservative 82; Mismatches 208; Indels 114; Gaps 19;

Db 39 GEMTSGONENRHSPLTOTITENVOGLQVWARGMOPK-VQVTPLIHGYMYLANP 97
41 GEMTHGDDAGTRISLAQITPPMAKELIWSYDLSSGVAETPIYVDGVMTAPW 100
98 DYIQAIDAKGDLIEMHRROLPIATLNS-FGIYANGYIV-AGSTCOYSPFGCVSGHDS 155
101 SYVHALDVSGSKRLMTIDPEVPREKGNACCDVNRGVAVHECKFVGSIDGRIVA-IDA 159
156 ANGEIM-RNFI-----PRA--GEEDETWGNVDYBAAMMGAM--GO 193
160 PGRKIVMERNTLIDDKFYTTGAPRYINGGAEVGRGITYATDPYASRPV 219
194 I-----TYDPTNLVHYGSTAVG----- 211
220 VPBGPSLPEDASMAAKTWDPAGGVGSGRRHGYELDGLYRKAGCCISAPAPPS 279
212 PASTGORTGGLTYNTNRFAVPPDGGELIVNRHOTLPROMWDECTFPMVNTVNDV-Q 289
280 PMSRRKSPAGDNLTYASTIYALRPTDGETVNHVQOTPRADNMWYSTODLLADIEIGK 339
270 PSTEMEGLOSINPMAAGERRVLTGVPCKTGTMMQDATGTFEIMARDTYONMIESID 341
340 P-----RKVIILHP-KNGFEFVIDRTDCKFTSKAONVFPYMAWATGYDE 380
330 NGIYTVN-EDALIKELDEYDVCEFTLGRDWPSSAALNPSGTYFPILNVCYIDMAVDO 388
381 NGRIENPAGAPGHLISMR--PAPSARTMHSMSYSPTGLAFPAONIP-L-VLOEDK 436
389 EFTSMYVNTSN-----VTLPPGKDMIGRIDALIDISTGRTLSVERAAA 433

RESULT 10
O9AP95 PRELIMINARY; PRT; 691 AA.
AC O9AP95
DT 01-DEC-2001 (TrEMBLrel. 17, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
GN Pseudomonas butanovora.
OS Pseudomonas butanovora.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
NCBI_TaxID=86174;
RP SEQUENCE FROM N.A.
RA Varginal A.S.; Alp D.J.; Sayavedra-Soto L.A.;
RT dehydrogenases involved in the expression of two distinct alcohol
KT butanovora.
RL submitted (NOV-2001) to the EMBL/GenBank/DBJ databases
DR EMBL; AF355798; AAK27220.2; Pseudomonas
SQ SEQUENCE 691 AA; 75070 MW; 4EC7FD20CDAA14E64 CRC64;

OY		265	NNDVDPSTEMEGLQSIINNAATGERRVLTGPCKTGTMMQFADTEGFILMA---	IDNTNQ	321
Db		348	E-----QTRKE-----GKKRLLTHPRKDNIVYTLDRENDDLLSADKLDTF--V		369
OY		322	NMIESIDENGIVTVNEDALIKELDVE-YDVCPFTLGGSDMPMSALNPDGSIGYPIPLNNVC		380
Db		390	NVFHVHDKLSGLPVRPEFGTRMDRKGEITCPSAGHYNCOSHDSYDPTKOLFPMGCJNHIC		449
OY		381	YDMAVAVOEFTSDMYNTSNVTCLP-PGKDM-----ICRIDALDISGRFLMSVERAAN		434
Db		450	MDMEPFMIPRAGEFYVGATLMYPPPKCDRONNYLGLOIAKAYNAVITYTKHEIMERSVS		509
OY		435	YSPLYSTGGGVLFNGSGTDRTFRALSOTGETTLMTROTIRLVASGSAISYEVDGMQYVLA		494
Db		510	WCGTLATAGNLVIFYGLTDGLFKLANSDTGELLKHKLPSGVIGYPMFEYHKGVQYVAMS		569
OY		495	G-GVSYGSGSL	503	
Db		570	GVGGMPPGYCL	579	
RESULT	12				
O9AGW3	PRELIMINARY;	PRT:	623 AA.		
ID	O9AGW3				
AC	01-JUN-2001 (TREMBLrel. 17, Created)				
DT	01-JUN-2001 (TREMBLrel. 17, Last sequence update)				
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)				
DE	1-BUTANOL DEHYDROGENASE BOH PRECURSOR.				
OS	Pseudomonas butanovora				
OC	Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;				
OC	Pseudomonas.				
OX	NCBI_Taxid=86174;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RA	Sayavedra-Soto L.A., Vangnai A.S., App D.J.;				
RT	"Characterization of the expression of two distinct alcohol				
RT	dehydrogenases involved in butane metabolism in Pseudomonas				
RL	butanovora."				
RL	Submitted (DEC-2000) to the EMBL/GenBank/DDBJ databases.				
DR	EMBL; AF326086; AKI5506.1; -				
DR	HSSP; Q9Z4J7; IFIG.				
DR	Interpro: IPR002372; Bac_PQQ_repeat.				
KW	Pfam: PF01011; Bacterial_PQQ_4.				
Signal.					
FT	1 29 POTENTIAL.				
SEQENCE	623 AA: 67553 MW: A573AlC64Ac12D55 CRC64;				
Query Match	15.6%; Score 422.5; DB 2; Length 623;				
Best Local Similarity	24.8%; Pred. No. 1.5e-22;				
Matches 157; Conservative	80; Mismatches 220; Indels 175; Gaps				
OY	2 KPTSLMASGALALLAAPAFACVTPYVDELIAN--PRAGMISYGONENYRHSPLOTI	59			
Db	7 KPFLRALIVATAAALSLPA-AAVTVTMEDLANDHKHTGVLTFGLGKAKORHSPLKAI	65			
OY	60 TTENVGQLQVMA-----RCMQGKVQVPLIHGVMYLAVGDIQDAKDGDLI	111			
Db	66 MTDNAANIYPAMSSPFGEKQRQGE---AQV-LVHDGIYATASYSTRIFALDARSOKRL	120			
OY	112 WEHRQLPNIATLNSFGIVANGVIYAGSTCOYSPFCFSGSDSATGELMLNRF----	166			
Db	121 WEYNLRLPD-DIRPCDQVNRGAALYDKVFFGTLDAAVAVALDRKTGVKVMKKRGDHRY	179			
OY	167 -----IPRAGEGDDTY-----	178			
Db	180 GYTMGAPRVINKDQSGRTLIVHGSSGDEFGVVGWLFAFPDTEGEVYNARPVBEGHMGL	239			
OY	179 -GND-----YEARMTGA--WCQITYPDVTNLVHYGSTAVPG	212			
Db	240 NGKDSIPTGDPKPAPSWPDDPNSPICKYDA-WSGGGA PMQTASFDEVNNNVYIGAAGNPAP	298			

```

OY 213 ASETORGTGPG-----GLTGTNTREAVRBDTGETIWMHROTLPDRNDMOECTEEMVTN- 265
DB 299 WNTWKRTPARPGDDRRNMDSLF-TSGOAYVDASTGELKGYOHTPRDADPEGNNNSVLEEX 357
OY 266 -----VDVOPSTEMEGLQSIIN-----PNAATGERVVLTCYCKTGTMMO-- 304
DB 358 KDRKRYGMVNASAHADRNGFFPVTRDMLAKAGCYPKKPS--LGAMFPEVDGLTWASG 414
OY 305 FQAEETGEFLWARDVYQNMIESIDENGIVYNEEDAILKELDERVDYDCTPELGRDMPSA 364
DB 415 FDLTKTKPI-EKDNRPQPKGADKGESIFVS-----PFLGCTGMHMS 458
OY 365 LNDSGIVPIPIANVCYDMMAVDOETSMVNTSNVT-----KLPG-----KMT 411
DB 459 YSPDTGLFYIPIANHWAMD-----TYTENVYTKKASATLGGCFRIKMLFDHV 505
OY 412 GIDAIDISTGRTLMSVERAANSPVLTSGGVLFNGSTDYRFRALSQETGETIMOTRL 471
DB 506 GILRAIDPSPARSLAGAGRVPA-VAGITLTAGGWFTGTSDBYLKAEPDAKNGKELMKFQT 564
OY 472 ATVASGOAISYEVDGMQYVALIAGGVSYSGSL 503
DB 565 GSGVYVPTWEMDGEQYVALIOG---YGVAV 593

RESULT 13
OY 09L935 PRELIMINARY; PRT; 599 AA.
AC 09L935;
DT 01-OCT-2000 (TREMblrel. 15, Created)
DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
DE METHANOL DEHYDROGENASE LARGE SUBUNIT.
GN MXAF.
OS Methylovorus sp. (strain SS1 / DSM 11726).
OC Bacteria; Proteobacteria; beta subdivision; Methylophilus group;
OC Methylovorus.
OX NCBI_TaxID=81683;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=SS1;
RC Kim Y.M.;
RT "Cloning and nucleotide sequence of mxaf gene of Methylovorus sp.
RT strain SS1 DSM11726."
RL Submitted (SEP-1999) to the EMBL/Genbank/DBJ databases.
DR EMBL; AF184915; AAD56237.2;
DR HSSP; P38539; 4AAH.
DR InterPro; IPR001479; Bac_PCO.
DR InterPro; IPR002372; Bac_PCO_repeat.
DR Pfam; PF01011; Bacterial_PCO; 7.
DR PROSITE; PS00364; Bacterial_PCO; 7.
SO SEQUENCE 599 AA; 65133 MW; DBF6F4B5D871BC91 CRC64;

```

```

Query Match 15.48; Score 417.5; DB 2; Length 599;
Best Local Similarity 24.48; Pred. No. 3.3e-22;
Matches 142; Conservative 97; Mismatches 229; Indels 115; Gaps 21;

```

```

OY 4 TSLWASGALALLAAPAFQVTPVDELANPAGEMISTYQONENRHSFLQITEN 63
DB 6 TATGAAVAGLALSMLPSVAADSLAALGAN--NNMTMQGDTTGGYHSRLSOLITGN 63
OY 64 VGOQL---WARMOPGKVOYVELLHDGVYLAN--PDYQIADKATGLLMEHR--RQ 117
DB 64 VKNLESGIGFSTGLTGH-EGAPLVIGDMYINTPFNNFALNALPEKIVYQKHPKO 122
OY 118 LPNATATNSGYANGV-----IVA-----GSTCY 143
DB 123 DASKAAVACDDIYNRGLAYDGHIFKTQLDGHIVASDAKTKGKELMKMNCDAVAGSTITQ 182
OY 144 SPF-----GC-----FVSGHDSATGEELMKNYFT-----PRA 170
DB 144 SPF-----GC-----FVSGHDSATGEELMKNYFT-----PRA 170

```

```

DB 183 APFAVAKVLVCGSGAELGVGTYATYDOKTGLVMSFATGPDNLANAKDFNKDNPHY 242
OY 171 GEE--GDETFWNGDYEARMTGA---WGOITDPTNLVHGSTANGAPASETGRTGTL 225
DB 243 GOMGLGKTKEGE---GKIGGTFNMGWYADYDKRLMLFYFGSGNAPNWTMR--PGDNK 297
OY 226 YGNTREFAVRPDTGEIYWRHQTLPDRNDMOECTEEMVTNVDPVOPSTEMEGLQSIINPAA 285
DB 298 W-TMTIARVDVTAAMKMGYOKTPHDEMDFAGVNMILLDOAVNCKTQ----- 344
OY 286 TGERVVLTCYCKTGTMMOFAETGEFLWARDN--YQNMIESIDENGIVYNEEDAILKEL 344
DB 345 ----PLTHVD--RNGIYLTNRQTSIVQAKVDVAVVEFKVDKTKGLPVRDEPESSTR 399
OY 345 DVE-YDVCEPTEFLGRDMPSAALNPDGIFITPLANVCYDMMAVDOETSMVNTSNVT 403
DB 400 DHKSTNICPSAMGFHNGIGLADVDPSRTTFYGLNHCIDMEPEMLPYRAGQFVGATLAM 459
OY 404 LP---PGKMTIGRIDAIDISTGRTLMSVERAANSPVLTSGGVLFNGSTDYRFRALS 459
DB 460 YPGPNQPTKEMGOYVLAIDGVYGEVKKWTKKFEFSVWGTLATGGLVFTNTLDGNIKALD 519
OY 460 QETGETIMOTRLATYASGOAISYEVDGMQYVALIAGGVSYSGS 502
DB 520 KTKGEIKKFKMPSGALIGAPMSYAKKQYIA-----TNYGVG 557

```

```

RESULT 14
OY 0934G0 PRELIMINARY; PRT; 695 AA.
AC 0934G0;
DT 01-DEC-2001 (TREMblrel. 19, Created)
DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)
DE LUPANINE HYDROXYLASE PRECURSOR.
GN LUD.
OS Pseudomonas sp. DH2001.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=163560;
RN [1]
RP SEQUENCE FROM N.A.
RA Hopper D.J., Kaderbhai M.A., Little A.R., Marriott S.A., Young M.,
RA Rogozinski J.;
RT "Cloning, sequencing and analysis of the gene for lupanine
RT hydroxylase, a quinuoylchome c from a Pseudomonas sp."
RL Submitted (JUN-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL; AJ318095; CAC67410.1;
DR KMW SIGNAL.
FT SIGNAL.
FT CHAIN 1 26 LUPANINE HYDROXYLASE.
SO SEQUENCE 695 AA; 74956 MW; 9096C6387E457FE0 CRC64;

```

```

Query Match 15.18; Score 408; DB 2; Length 695;
Best Local Similarity 25.88; Pred. No. 2e-21;
Matches 139; Conservative 71; Mismatches 218; Indels 110; Gaps 19;

```

```

OY 38 AGENISTGQNOENYRHSPLQITTEVNGQLOLYMARQGPCKOV--TPLIHDGVYLANP 96
DB 36 SGWMSLLGGGNEQHYFSALKDVNKSVAKNLGLSWPTDMEAGDGLVGNPLVAQDVIRGGP 95
OY 97 GDVIOAIDAKTGLLMEHRQLPNATLNSFGI---VANGVY-----ASTGQ----- 142
DB 96 PKIIVYANDLTKGNLMTYTYPEVOYDKTSWTGFGCTHVNGGLAVDDNNYISGYCKLLAV 155
OY 143 -----YSPFGC-----FVS-----GH-----DSATGEEL 161
DB 156 SRTTRKILWSSQSCOPKMKQAITGAPRYGCKVFIQNASGDFGCDRGHDAFADATKCHL 215
OY 162 WRNYELPRAGE-----EGDETWNNDYEARMM-----TGAMGQITVDPTVNLVHY 205
DB 216 WRFTYMP--GDSKRFENDLLAKASKITWGDY---WKYIKGCVSPWDAITTYDEASDTLYP 270

```

```

QY  AGGEMISYGQNGENYRHSPLTQTTENYVOLDIW--AAGMPGAYQVTPLIHGGVMY---92
      || : : ||| || : : || : : || : : || : : || : : || : : || : : ||
Db  AGAMPATGCTGYTSHNSPLAQINKSNVAKYKNAANSFSTGLNGH--EGAPLYIDMMYVHS  71

QY  -----LANPGDVI-----QA 1021
      : || : :
Db  72 APFNNTYALNLMDPGKTIWQHFKQDASTKAVMCCDDVDRLAGAGQIVKQANGHLLA  1311

QY  IDAKTGDLIWEHRQLDPI--ATLNSFGIYANGVYASGTQYSPFGC--PFSCHDSATGE  159
      : ||| : : ||| : : ||| : : ||| : : ||| : : ||| : : ||| : : |||
Db  132 LDKAKTGKINWEVCDPKKSGSLTGAFAFAKDTYLMG--CSGAELGRGAVNAADLTGE  1889

QY  ELMR-----NYFIPRAG--EGDETQNDCEYARMWTA---WQOITYDP  1989
      || : : || : : || : : || : : || : : || : : || : : || : : ||
Db  190 LKMAAPETGSDSDVRLAKPQNSANPBYGQFGTGTWEGD---AMKIGGCTNMGWYAVDP  2466

```

Search completed: May 24, 2002, 10:30:19
Job time: 652 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2002, 10:19:20 ; Search time 101.42 Seconds
(without alignments)
560.735 Million cell updates/sec

Title: WALICK-934-135.PEP
Sequence: 1 MKPSTLWASAGALALLAP.....GMOYVAGGVSYSGSLNS 512

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues
Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A-Geneseq_032802.*
1: /SIDSL/gcgdata/hold-geneseq/genesep-emb1/AA1980.DAT.*
2: /SIDSL/gcgdata/hold-geneseq/genesep-emb1/AA1981.DAT.*
3: /SIDSL/gcgdata/hold-geneseq/genesep-emb1/AA1982.DAT.*
4: /SIDSL/gcgdata/hold-geneseq/genesep-emb1/AA1983.DAT.*
5: /SIDSL/gcgdata/hold-geneseq/genesep-emb1/AA1984.DAT.*
6: /SIDSL/gcgdata/hold-geneseq/genesep-emb1/AA1985.DAT.*
7: /SIDSL/gcgdata/hold-geneseq/genesep-emb1/AA1986.DAT.*
8: /SIDSL/gcgdata/hold-geneseq/genesep-emb1/AA1987.DAT.*
9: /SIDSL/gcgdata/hold-geneseq/genesep-emb1/AA1988.DAT.*
10: /SIDSL/gcgdata/hold-geneseq/genesep-emb1/AA1989.DAT.*
11: /SIDSL/gcgdata/hold-geneseq/genesep-emb1/AA1990.DAT.*
12: /SIDSL/gcgdata/hold-geneseq/genesep-emb1/AA1991.DAT.*
13: /SIDSL/gcgdata/hold-geneseq/genesep-emb1/AA1992.DAT.*
14: /SIDSL/gcgdata/hold-geneseq/genesep-emb1/AA1993.DAT.*
15: /SIDSL/gcgdata/hold-geneseq/genesep-emb1/AA1994.DAT.*
16: /SIDSL/gcgdata/hold-geneseq/genesep-emb1/AA1995.DAT.*
17: /SIDSL/gcgdata/hold-geneseq/genesep-emb1/AA1996.DAT.*
18: /SIDSL/gcgdata/hold-geneseq/genesep-emb1/AA1997.DAT.*
19: /SIDSL/gcgdata/hold-geneseq/genesep-emb1/AA1998.DAT.*
20: /SIDSL/gcgdata/hold-geneseq/genesep-emb1/AA1999.DAT.*
21: /SIDSL/gcgdata/hold-geneseq/genesep-emb1/AA2000.DAT.*
22: /SIDSL/gcgdata/hold-geneseq/genesep-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2710	98.8	579	19 AAW37873	Alcohol and/or ald
2	2385	87.0	579	19 AAW37874	Alcohol and/or ald
3	2328.5	84.9	579	19 AAW37876	Alcohol and/or ald
4	2291.5	83.6	578	19 AAW37875	Alcohol and/or ald
5	484	17.7	754	21 AAB35987	ADH complex protei
6	483.5	17.6	738	13 AAR20192	A. allicetigenes me
7	482.5	17.6	738	11 AAR13993	A. allicetigenes me
8	482	17.6	742	11 AAR05235	Amino acid sequenc
9	315.5	11.5	740	20 AAW95019	Sorbitol dehydroge
10	158.5	5.8	443	22 ABG24430	Novel human diagno
11	146	5.3	824	18 AAW09614	Pathogenic Scaphy1

12	144.5	5.3	1016	22 AAW34349	Staphylococcus aur
13	134.5	4.9	715	22 AAG81738	S. epidermidis ope
14	134	4.9	910	22 AAG83007	S. epidermidis ope
15	133.5	4.9	696	22 AAG82914	S. epidermidis ope
16	130.5	4.8	470	22 AAG82701	Novel human diagno
17	129	4.7	948	22 AAG25909	Novel human diagno
18	129	4.7	1510	22 AAG21573	S. epidermidis ope
19	116.5	4.2	404	22 AAG82772	erya region polype
20	116	4.2	3567	14 AAR4431	Cryptosporidium pa
21	113	4.1	1042	21 AAB11732	Cryptosporidium pa
22	113	4.1	1837	21 AAB11726	Amino acid sequenc
23	112	4.1	790	22 AAG67230	Protein G variant
24	111.5	4.1	593	11 AAR07014	Sequence of polype
25	110.5	4.0	480	8 AAB70468	Protein G variant
26	110	4.0	649	22 AAB66538	Drosophila melanog
27	108.5	4.0	530	22 AAW10236	Human PRO4340 poly
28	107.5	3.9	501	18 AAW10645	HPV6 mutant L1 pro
29	107	3.9	269	11 AAR07006	Protein G variant.
30	107	3.9	269	15 AAR53295	IgG-binding Strept
31	106.5	3.9	501	18 AAW10644	Human protein sequ
32	106.5	3.9	514	22 AAB95225	Human prostate can
33	106.5	3.9	594	12 AAR10005	S. epidermidis ope
34	106	3.9	269	10 AAB94788	C. trachomatis I2
35	106	3.9	1012	20 AAY16735	Cryptosporidium pa
36	105.5	3.8	429	22 AAG66542	Drosophila melanog
37	105.5	3.8	429	22 AAB93799	Human Interferon-a
38	105.5	3.8	469	21 AAB56578	Human protein sequ
39	105	3.8	278	22 AAG82625	S. epidermidis ope
40	105	3.8	505	20 AAT16739	C. trachomatis I2
41	104.5	3.8	1721	19 AAB48299	Cryptosporidium pa
42	104	3.8	1429	22 AAB58779	Drosophila melanog
43	103.5	3.8	501	18 AAW26108	HPV11 mutant L1 pr
44	103.5	3.8	501	18 AAW26109	HPV11 mutant L1 pr
45	103.5	3.8	593	15 AAR62944	Streptococcus Prot

ALIGNMENTS

RESULT 1	
AAW37873	AAW37873 standard; Protein: 579 AA.
XX	XX
AC	AAW37873;
DT	10-AUG-1998 (first entry)
XX	XX
DE	Alcohol and/or aldehyde dehydrogenase A amino acid sequence.
XX	XX
KW	Alcohol/aldehyde dehydrogenase A enzyme; recombinant organism;
KW	aldehyde; ketone; carboxylic acid; L-sorbose; D-sorbitol;
KW	2-keto-L-gulononic acid; L-ascorbic; inhibition.
XX	XX
OS	Gluconobacter oxydans.
XX	XX
FH	Key
FT	Peptide
FT	Location/Qualifiers
FT	1..23
FT	/note="signal peptide"
FT	24..579
FT	/note="mature protein"
PN	EP832974-A2.
XX	XX
PD	01-APR-1998.
XX	XX
PF	11-SEP-1997; 97EP-0115801.
XX	XX
PR	19-SEP-1996; 96EP-0115001.
XX	XX
PA	(HOFF) HOFFMANN LA ROCHE & CO AG F.
XX	XX
PI	Asakura A, Hoshino T, Ojima S, Shinjoh M, Tomiyama N;
XX	XX

DR WPI: 1998-195228/18.
N-PSDB: AAV29051.

Recombinant Gluconobacter oxydans alcohol and/or aldehyde dehydrogenase enzyme(s) - useful for converting substrate(s), e.g. L-sorbose or D-sorbitol to 2-keto-L-gulonic acid

Claim 1: Pages 35-37; 59pp: English.

This is the amino acid sequence for the Gluconobacter oxydans alcohol and/or aldehyde dehydrogenase A enzyme. The enzymes or recombinant organisms can be used to convert suitable substrates to aldehydes, ketones or carboxylic acids, especially to convert L-sorbose or D-sorbitol to 2-keto-L-gulonic acid, which can be converted to L-ascorbic acid by standard procedures. The derivatives of ADH enzymes have desired substrate specificity, higher affinity to a substrate, lower affinity to an inhibitory compound, higher stability against temperature and/or pH and higher catalytic speed.

Sequence 579 AA:

Query Match 98.8%; Score 2710; DB 19; Length 579;
Best Local Similarity 92.1%; Pred. No. 2.3e-225;
Matches 512; Conservative 0; Mismatches 0; Indels 44; Gaps 1;

```

OY 1 MKPTSLMASGALALLAAPAFVPTDELTPANPAGEMI SYGONONRYHSPLTQIT 60
DB 1 mkptslmasgalaallapafvptdeltpanpagemi sygononryhspltdit 60
OY 61 TENVGQLOLVARMGOPKRVQVTPLIHDGVMTIANPGVIOAIDAKGDLWEHRQLPN 120
DB 61 tenvgqlolvarmgopkrvqvtplihgdvmytlanpgvviqaidaktgdlwehrrqlpn 120
OY 121 IATLNSFGEPTRGMA-----I 136
DB 121 iatlnsfgeptrgmalylgtnyfvswdnhlvaldcatgvtfdvdrqgedmvsnsqpi 180
OY 137 VANGVIVAGSTCOYSPFGCVSGHDSATGELMRYFTIPRAGEGDETMGNDYEARMTG 196
DB 137 vangvivaagstcoyspfgcvsgshdsatgelmryftipragegdetmgnndyearmtg 196
OY 181 VANGVIVAGSTCOYSPFGCVSGHDSATGELMRYFTIPRAGEGDETMGNDYEARMTG 240
DB 181 vangvivaagstcoyspfgcvsgshdsatgelmryftipragegdetmgnndyearmtg 240
OY 197 AMGOITVDPVTNLVHGSTAVGPASETORPGTILYGTNRAPVPPDGETIWRHQTLP 256
DB 197 amgoitvdpvtnlvhgstavgpasetorgptilgtntnravppdgetiwrhqtlp 256
OY 241 AAGQGLYDVPVTNLVHGSTAVGPASETORPGTILYGTNRAPVPPDGETIWRHQTLP 300
DB 241 aagqglydvpvtnlvhgstavgpasetorgptilgtntnravppdgetiwrhqtlp 300
OY 257 RDNMDOCTFEMMYTNNVDOVOSTEMEGLOSINMATGERRVLGVPCKTGMWQFDAET 316
DB 257 rdnmdoctfemmytnnvdoovostemeglosinmatgerrvlgvpcktgmwqfdaet 360
OY 301 RDNMDOCTFEMMYTNNVDOVOSTEMEGLOSINMATGERRVLGVPCKTGMWQFDAET 360
DB 301 rdnmdoctfemmytnnvdoovostemeglosinmatgerrvlgvpcktgmwqfdaet 360
OY 317 GEFIMARDTNTONMIESIDENGIWVEDALIKELDEYDVOPFTFLGGDMPSALNDPS 376
DB 317 gefimardtntonmiesidengiwvedalikeldeydvopftflggdmpsalndps 420
OY 377 GIVFIPINNYCYDMMAVDQETSMOYNTSNVTKLPPGKDMIGRIDAIDISTGRTLSMVE 436
DB 377 givfipinnycydmmavdqetsmoyntsnvtklppgkdmigridaidistgtrtlsme 480
OY 421 GIVFIPINNYCYDMMAVDQETSMOYNTSNVTKLPPGKDMIGRIDAIDISTGRTLSMVE 480
DB 421 givfipinnycydmmavdqetsmoyntsnvtklppgkdmigridaidistgtrtlsme 480
OY 437 RAAANSPVLSTGCVLFNGTIDRYFRALSOETGELMOTRLATYASQALSYEVDMQY 496
DB 437 raaanspvlstgcvlfngtidryfralsetgelmotrlatyasqalsyevdmqy 496
OY 481 RAAANSPVLSTGCVLFNGTIDRYFRALSOETGELMOTRLATYASQALSYEVDMQY 540
DB 481 raaanspvlstgcvlfngtidryfralsetgelmotrlatyasqalsyevdmqy 540
OY 497 VALAGGVSYSGSLNS 512
DB 497 valaggvysygslns 556
OY 541 VALAGGVSYSGSLNS 556
DB 541 valaggvysygslns 556

```

RESULT 2
AAW37874
ID AAW37874 standard: Protein: 579 AA.
AC AAW37874;
XX

DT 10-AUG-1998 (first entry)

XX Alcohol and/or aldehyde dehydrogenase A' amino acid sequence.

DE Alcohol/aldehyde dehydrogenase A' enzyme; recombinant organism;

XX aldehyde; ketone; carboxylic acid; L-sorbose; D-sorbitol;

KW 2-keto-L-gulonic acid; L-ascorbic; inhibition.

XX Gluconobacter oxydans.

OS Key Location/Qualifiers

XX Peptide 1..23

FT Protein /note= "signal peptide"

FT Protein 24..579

XX Protein /note= "mature protein"

PN EP832974-A2.

XX 01-APR-1998.

XX 11-SEP-1997; 97EP-0115801.

XX 19-SEP-1996; 96EP-0115001.

XX (HOFF) HOFFMANN IA ROCHE & CO AG F.

PI Asakura A, Hoshino T, Ojima S, Shinjoh M, Tomiyama N;

DR WPI: 1998-195228/18.

XX N-PSDB: AAV29052.

XX Recombinant Gluconobacter oxydans alcohol and/or aldehyde

PT dehydrogenase enzyme(s) - useful for converting substrate(s), e.g.

PT L-sorbose or D-sorbitol to 2-keto-L-gulonic acid

XX Claim 1: Pages 38-40; 59pp: English.

PS This is the amino acid sequence for the Gluconobacter oxydans

XX alcohol and/or aldehyde dehydrogenase A' enzyme. The enzymes

CC or recombinant organisms can be used to convert suitable substrates

CC to aldehydes, ketones or carboxylic acids, especially to convert

CC L-sorbose or D-sorbitol to 2-keto-L-gulonic acid, which can be

CC converted to L-ascorbic acid by standard procedures. The derivatives

CC of ADH enzymes have desired substrate specificity, higher affinity

CC to a substrate, lower affinity to an inhibitory compound, higher

CC stability against temperature and/or pH and higher catalytic speed.

XX Sequence 579 AA:

Query Match 87.0%; Score 2385; DB 19; Length 579;
Best Local Similarity 80.1%; Pred. No. 2.9e-197;
Matches 442; Conservative 32; Mismatches 34; Indels 44; Gaps 1;

```

OY 1 MKPTSLMASGALALLAAPAFVPTDELTPANPAGEMI SYGONONRYHSPLTQIT 60
DB 1 mkptslmasgalaallapafvptdeltpanpagemi sygononryhspltdit 60
OY 61 TENVGQLOLVARMGOPKRVQVTPLIHDGVMTIANPGVIOAIDAKGDLWEHRQLPN 120
DB 61 tenvgqlolvarmgopkrvqvtplihgdvmytlanpgvviqaidaktgdlwehrrqlpn 120
OY 121 IATLNSFGEPTRGMA-----I 136
DB 121 iatlnsfgeptrgmalylgtnyfvswdnhlvaldcatgvtfdvdrqgedmvsnsqpi 180
OY 137 VANGVIVAGSTCOYSPFGCVSGHDSATGELMRYFTIPRAGEGDETMGNDYEARMTG 196
DB 137 vangvivaagstcoyspfgcvsgshdsatgelmryftipragegdetmgnndyearmtg 196
OY 181 VANGVIVAGSTCOYSPFGCVSGHDSATGELMRYFTIPRAGEGDETMGNDYEARMTG 240
DB 181 vangvivaagstcoyspfgcvsgshdsatgelmryftipragegdetmgnndyearmtg 240
OY 197 AMGOITVDPVTNLVHGSTAVGPASETORPGTILYGTNRAPVPPDGETIWRHQTLP 256
DB 197 amgoitvdpvtnlvhgstavgpasetorgptilgtntnravppdgetiwrhqtlp 256
OY 241 AAGQGLYDVPVTNLVHGSTAVGPASETORPGTILYGTNRAPVPPDGETIWRHQTLP 300
DB 241 aagqglydvpvtnlvhgstavgpasetorgptilgtntnravppdgetiwrhqtlp 300

```


XX 01-APR-1998.
 PD 11-SEP-1997; 97EP-0115801.
 XX 19-SEP-1996; 96EP-0115001.
 XX (HOFF) HOFFMANN LA ROCHE & CO AG F.
 PA Asakura A, Hoshino T, Ojima S, Shinjoh M, Tomiyama N,
 PI WPI: 1998-195228/18.
 DR N-PSDB; AAV29053.
 XX
 PT Recombinant Gluconobacter oxydans alcohol and/or aldehyde
 PT dehydrogenase enzyme(s) - useful for converting substrate(s), e.g.
 PT L-sorbose or D-sorbitol to 2-keto-L-gulononic acid
 PS Claim 1; Pages 41-43; 59pp; English.
 XX
 CC This is the amino acid sequence for the Gluconobacter oxydans
 CC alcohol and/or aldehyde dehydrogenase A' enzyme. The enzymes
 CC to recombinant organisms can be used to convert suitable substrates
 CC to aldehydes, ketones or carboxylic acids, especially to convert
 CC L-sorbose or D-sorbitol to 2-keto-L-gulononic acid, which can be
 CC converted to L-ascorbic acid by standard procedures. The derivatives
 CC of AADH enzymes have desired substrate specificity, higher affinity
 CC to a substrate, lower affinity to an inhibitory compound, higher
 CC stability against temperature and/or pH and higher catalytic speed.
 XX
 SO Sequence 578 AA;

Query Match 83.6%; Score 2291.5; DB 19; Length 578;
 Best Local Similarity 76.9%; Pred. No. 3.4e-189;
 Matches 427; Conservative 37; Mismatches 46; Indels 45; Gaps 2;

OY 1 MKPTSLMAASGALALIAAPAAQVTPVDELAMPAGEMSYGONENRHSPLQIT 60
 DB 1 MKTLTSSAALLVGLPALAQ-TALIDEMLANPPGGEWYNGQENYRHSPLQIT 59
 OY 61 TENVGOLQVWARGMOPGKQVTPPLIHGCVMLANPGVIOAIDAKTGDLTWEHRRLPN 120
 DB 60 ADVNGQLVWARGMEAGKIVPLVHDGVMLANPGVIOAIDAKTGDLTWEHRRLPN 119
 OY 121 IATLNSFGEPTRGMA-----1 136
 DB 120 IATLNSFGEPTRGMA-----1 136
 OY 137 VANGVIVASTCOYSPFGCFVSGHDSATGEELMKNYFIPRAGEGDETTWGNDEYARMNG 196
 DB 180 VANGVIVASTCOYSPFGCFVSGHDSATGEELMKNYFIPRAGEGDETTWGNDEYARMNG 196
 OY 197 AAGGOTTTPVNLVHGSTAVPASSETORGPGLTGTMTARAVRPPDGEYWRKQTL 256
 DB 240 VVGQTLTPVNLVHGSTAVPASSETORGPGLTGTMTARAVRPPDGEYWRKQTL 256
 OY 257 RDNMQOCTEFMAYTNDVOPSTMEGLSINPAATGERSRLTGVPCKTGTMOFDAET 316
 DB 300 RDNMQOCTEFMAYTNDVOPSTMEGLSINPAATGERSRLTGVPCKTGTMOFDAET 316
 OY 317 GEELMARNYNNQNIIESIDENGIYVNEALTELKELDERDVOPFTLGGDMWSALANDS 376
 DB 360 GEELMARNYNNQNIIESIDENGIYVNEALTELKELDERDVOPFTLGGDMWSALANDS 376
 OY 377 GIVEFIPINACYSMTAVDEFTSMYNTKLPKGMIGRIDADISTGRTLSVE 436
 DB 420 GIVEFIPINACYSMTAVDEFTSMYNTKLPKGMIGRIDADISTGRTLSVE 436
 OY 437 RAAANTSPVLSTGGVLFNGGDRFRALSOPTGETLWOTLATVYASQALSTVEDMOY 496
 DB 480 RYANAPVLTSGVLFNGGDRFRALSOPTGETLWOTLATVYASQALSTVEDMOY 496

OY 497 VAIAGGVSYSGSLN 511
 DB 540 VALGRGTSYSGSLN 554

RESULT 5
 AAB35987
 ID AAB35987 standard; Protein; 754 AA.
 AC AAB35987;
 XX
 DF 01-MAR-2001 (first entry)
 XX
 DE Sorbitol dehydrogenase subunit 1 amino acid sequence.
 XX
 KW Sorbitol dehydrogenase; SDH; Gluconobacter oxydans; D-sorbitol;
 XX L-sorbose production; 2-keto-L-gulononic acid.
 XX
 OS Gluconobacter oxydans.
 XX
 PN WO200065066-A1.
 XX
 PD 02-NOV-2000.
 XX
 PF 23-APR-1999; 99WO-1B00736.
 XX
 PR 23-APR-1999; 99WO-1B00736.
 XX
 PA (CHOI/) CHOI E.
 PA (RHEE/) RHEE S.
 PA (LEEE/) LEE E.
 XX
 PI Choi E, Rhee S, Lee E;
 XX
 DR WPI: 2000-687351/67.
 DR N-PSDB; AAC83153.

PT Novel membrane-bound sorbitol dehydrogenase nucleic acid molecule
 PT isolated from Gluconobacter suboxydans useful for the fermentative
 PT production of 2-keto-L-gulononic acid and L-sorbose from D-sorbitol
 XX
 PS Claim 1; Fig 8; 96pp; English.
 XX
 CC This invention relates to an isolated membrane-bound sorbitol
 CC dehydrogenase (SDH) from Gluconobacter oxydans. The invention includes
 CC polynucleotides (AAC83153 - AAC83154) encoding the 3 subunits of SDH
 CC (AAB35987 - AAB35989). Also included in the invention are two
 CC polynucleotide sequences AAC83156 and AAC83157 which contain the subunit
 CC coding regions. The sorbitol dehydrogenase polynucleotide sequences are
 CC useful for producing L-sorbose from D-sorbitol and for increasing the
 CC production of 2-keto-L-gulononic acid by transforming a host cell,
 CC especially Gluconobacter with the DNA and selecting the transformed host
 CC cell. Note: An additional 66 DNA sequences (fragments of the SDH subunit
 CC genes) encoding fragments of SDH are specifically claimed, however these
 CC sequences are not given separately in the specification but are included
 CC in sequences AAC83156 and AAC83157.
 XX
 SO Sequence 754 AA;

Query Match 17.7%; Score 484; DB 21; Length 754;
 Best Local Similarity 25.9%; Pred. No. 8e-33;
 Matches 156; Conservative 81; Mismatches 239; Indels 126; Gaps 16;

OY 14 LALLAFAFAVTPV-----TDELLANPAGENISYGOENRHSPLQITTEWNGQ 66
 DB 17 LQCAALACATSPVALAEDTGTALINADHPDMSYGRYSRTPSLDQITKDNASN 76
 OY 67 LQVWARGMOPGKQV-TPLIHGCVMLANPGDVIOAIDAKTGDLTWEHRRLPN-NIATL 124
 DB 77 LKLEWYDLDNTRGGEPLVGVYMACTIVSKMALDAAGKLIWSYDQVPGNADR 136
 OY 125 NSFGEPTRGMAI-----136

Sequence 443 AA;

	Indels	Gaps	7;
6 LWAS-----ACGATATTAAAGTGGTTCTGTTA	49;		
QY			

302 vs 303

RESULT	11
AAW09614	
ID	AAW09614 standard; protein; 824 AA
XX	
AC	AAW09614;
XX	
DT	20-MAY-1997 (first entry)
XX	

P1	Nucleic acids specific to pathogenic Staphylococcus aureus strains -
PT	for diagnosis and immunogenic protein prodn.
XX	

[illegible]

-----AWGQITYDPVTNLVHYGSTAVGPA----- 220

200 ----- NPNAATGERVLTGPCKTGTMMQFDAQETGEELWARDTNYQNMIE 332

```
733 eygetieiprhdeffbkjplteke---evpgkygik--nptgtgv-----irpvd
|::|||
333 SIDENGIVTNEADILKELDYEDVDCPEFLGGRDPSALNP 375
|::|||
```

Db 780 svlykg--pvkgdsivekeelpfe-----kerkimpd 809

RESULT 12

AA034349 standard: Protein: 1016 AA.

AA034349:

14-FEB-2002 (first entry)

Staphylococcus aureus cellular proliferation protein #625.

Antisense: prokaryotic cellular proliferation protein;

antibiotic; antibacterial; drug design.

Staphylococcus aureus.

WO200170955-A2.

27-SEP-2001.

21-MAR-2001; 2001WO-US09180.

21-MAR-2000; 2000US-191078P.

23-MAY-2000; 2000US-206848P.

26-MAY-2000; 2000US-207727P.

23-OCT-2000; 2000US-242578P.

27-NOV-2000; 2000US-253625P.

22-DEC-2000; 2000US-257931P.

16-FEB-2001; 2001US-269308P.

(ELIN-) ELITRA PHARM INC.

Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;

Yamamoto RT, Xu HH;

WPI: 2001-611495/70.

N-PSDB: AAS52208.

New polynucleotides for the identification and development of

antibiotics, comprise sequences of antisense nucleic acids -

Example 3: Seq ID No 5845; 51pp; English.

The invention relates to antisense inhibitors of genes essential to

prokaryotic cellular proliferation, their use in identifying the

genes, their use in the discovery of novel antibiotics, the essential

genes themselves and the encoded proteins. The prokaryotes used are

Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella

pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The

invention is also useful for the identification of potential new targets

for antibiotic development. The antisense nucleic acids can also be used

to identify proteins used in proliferation, to express these proteins,

and to obtain antibodies capable of binding to the expressed proteins.

The proteins can be used to screen compounds in rational drug discovery

programmes. The antisense nucleic acid sequence is also useful to screen

for homologous nucleic acids which are required for cell proliferation in

essential prokaryotic cellular proliferation protein.

Note: The sequence data for this patent did not form part

of the printed specification, but was obtained in electronic

format directly from WIPO at

ftp.wipo.int/pub/published_pcl_sequences.

Sequence 1016 AA;

Query Match 5.3%; Score 144.5; DB 22; Length 1016;

Best Local Similarity 21.3%; Pred. No. 0.0027;

Matches 115; Conservative 63; Mismatches 170; Indels 191; Gaps 28;

19 AFAAQT-----FVTELLANPPAGEMISYCONQENYRHSPLTOTTENNQOLQV 70

Db 528 apgtkvtrregqgkgtlttptcknplgvliskgeskeetkdpinelt----- 577

QY 71 WAROMPGKQVQVPTLIDH-----GYMVLANGCDVLOAIDAKTGDLINRHR 115

Db 578 -----eygpeltpghndefqkplptgekeevpgkpkpkn-----etgadv----- 619

QY 116 ROLPNITLNSFGEPGMAIVANGYIVAGSTCOYSPFCFVSGHDSATGELMRNVRFP 175

Db 620 --rppvdsvlkyg-pvkgdsivekeelpfekeekfn-----dlapylek-----vt 663

QY 176 RAGEGDETMGNDYEAARMTGA-----WGQTTDPVNTLVHGSTAVGPA----- 220

Db 664 regqgkeltltptcknplgvliskgepkeetkdpineltteygpeltpghndefqk 723

QY 221 -----SEFQRTPGTGLTGTTRFAVRPDT-----GEIIVRHQTLPRDNDQECT 265

Db 724 lptgekeevpgkpg--lknpetgdvvrpydsvlkygvpvkgdsivekeelp--lkeek 778

QY 266 FEMMYTNVDVOPSTEMEGLSINPMATGERRVLTGVPCKTGMQGFDAETGEFLWARDT 325

Db 779 f-----nplapgte-----kvtrregqgkeltlt--plcknpl-----tgeil----- 815

QY 326 NQNMTESTIDENGIVTVEDATIKELDVEYDVCPFLUGRDWPSAL----- 372

Db 816 ---skgesiee-----itkdpinelteygpeltpghndefqkplptgekeevpgkpg 865

QY 373 --NPDGIVFIFPLNNVCDMAVDQFTSMVYNTKLPCK--DMIGRIDAIDISTG 429

Db 866 lknpetgdvvrp-----pvd-----svkygvpvkgdsivekeelpfeke 904

QY 430 RTLMSVRAAANSPVLSGTGGVLFNGGTDYFRALSGFTGTLWQRLATVAGSGAIS 488

Db 905 f-----kimpalap-----gtekvlre-gkgekeltpcknplgvlis 944

RESULT 13

AA081738 standard: Protein: 715 AA.

AA081738:

03-SEP-2001 (first entry)

S. epidermidis open reading frame protein sequence SEQ ID NO:570.

Staphylococcus epidermidis sri strain; infection; diagnosis;

Staphylococcus epidermidis.

WO200134809-A2.

17-MAY-2001.

09-NOV-2000; 2000WO-US30782.

09-NOV-1999; 99US-0164258.

(GLAX) GLAXO GROUP LTD.

Kimmerly WJ;

WPI: 2001-316495/33.

N-PSDB: AAS52588.

Nucleic acids encoding polypeptides from Staphylococcus epidermidis,

useful for vaccinating against infections, e.g. endocarditis -

Claim 18; Page 189; 2188pp; English.

AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides

(II), given in AAG81454 to AAG83120, from Staphylococcus epidermidis.

CC (I) and (II) can have antibacterial activity and therefore can be used
 CC in vaccination. The nucleic acids (I) may be used to produce the
 CC S. epidermidis polypeptides (II) via the production of vectors
 CC containing them which are used to produce hosts cells which express the
 CC polypeptides. The polypeptides (II) (and/or nucleic acids) may then be
 CC used to vaccinate subjects and to raise antibodies against the bacteria.
 CC The polypeptides may also be used to assay for other inhibitors of their
 CC activity and therefore identify compounds that may be used for the
 CC treatment of S. epidermidis infections, e.g. endocarditis. AAH53971 to
 CC AAH5090 represent specifically claimed S. epidermidis genomic DNA
 CC AAH5098 represent oligonucleotide sequences and primers which are used
 CC in the exemplification of the present invention. AAH5091 to
 CC N.B. The present invention specifically claims all the polynucleotide
 CC sequences given in the sequence listing of the present specification,
 CC however the sequence listing only goes up to SEQ ID NO:4454 so even
 CC though sequences are given in the disclosure for SEQ ID NO:4465 to 4472,
 CC no sequences are present for SEQ ID NO:4455 to 4464.
 CC
 CC Sequence 715 AA:

Query Match 4.9%; Score 134.5; DB 22; Length 715;
 Best Local Similarity 23.2%; Pred. No. 0.012; Mismatches 156; Indels 141; Gaps 26;

Matches 105; Conservative 51; Mismatches 156; Indels 141; Gaps 26;

26 TPVTDELLANPPAGEMISYQONENRHSPLQIT-----TENVOLOLWARQMO-76
 309 tptt-----knpltekgvgepektekvkqvdeivhyggeelkpgkdeifpnapsqg 364
 77 --PKR-----VOVPLIHGVMYLANPGDVIADAKGDLIMHRQL-PNIA--122
 365 dvppkpgvknptdtegvvtpvddvltkypvgdpl-----tsteepfdkkrefdpn 420
 123 --TINSGEPTRGMAIVANGVIVAGSTCOYSPFGCSGHSATGELMWRNYFIRAGEE 180
 421 tekvvkgep-----gtktitp-----tknpltekgv-----ge 451
 181 GDETWGNDYFARMGAMQITVDPYTNLHYGSTAVG-----ASENQROGT 227
 452 gept-----ekvtkqvdeivhyggeelkpgkdeifpnapsqgqtkp 497
 228 PGSTLYGTNTRRAVRPDTEGIYWRHQTLPRDNMOECTFEMAYTN--VDVOPSTMEGL 284
 498 pg-----vknpdtegvv-----tpvdd-----vtkygvdgdpitsteel 533
 285 -----OSINPMNATGERRY-LTGP-CKT-GTMMQFADATGEFLMARDNNGMIESIDE 336
 534 pfdkkrefnpldkpgeervkqkgpgetkltitpknpltekgvgepekteitkqvde 593
 337 NGITVNDAILKELDYEDVCPFLGGRWPS--AALNPDG-IVFIPLNWCV-----388
 594 --lteygggeelkpgkdeifpnapsqgedvppkpgvknptdtegvvtpvddvltkypvd 651
 389 -DNMAVNDG-FTSMADYNTSNVTKLPPGKDM 418
 652 gdpisteeelpfdkkrefnpl-----lpgkerv 680

RESULT 14

AAG83007
 ID AAG83007 standard; Protein; 910 AA.

AC AAG83007;

DT 03-SEP-2001 (first entry)

DE S. epidermidis open reading frame protein sequence SEQ ID NO:3108.

KW Staphylococcus epidermidis SRI strain; infection; diagnosis;

OS Staphylococcus epidermidis.

XX
 PN W0200134809-A2.
 XX
 PD 17-MAY-2001.
 XX
 PF 09-NOV-2000; 2000WO-0530782.
 XX
 FR 09-NOV-1999; 99US-0164258.
 XX
 PA (GLAX) GLAXO GROUP LTD.
 PI Kimmery MJ;
 DR WPI: 2001-316495/33.
 DR N-PSDB: AAH53857.
 XX
 PT Nucleic acids encoding polypeptides from Staphylococcus epidermidis,
 PT useful for vaccinating against infections, e.g. endocarditis -
 XX
 PS Claim 18; Page 819; 2188pp; English.

CC AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides
 CC (II), given in AAG81454 to AAG83120, from Staphylococcus epidermidis.
 CC (I) and (II) can have antibacterial activity and therefore can be used
 CC in vaccination. The nucleic acids (I) may be used to produce the
 CC S. epidermidis polypeptides (II) via the production of vectors
 CC containing them which are used to produce hosts cells which express the
 CC polypeptides. The polypeptides (II) (and/or nucleic acids) may then be
 CC used to vaccinate subjects and to raise antibodies against the bacteria.
 CC The polypeptides may also be used to assay for other inhibitors of their
 CC activity and therefore identify compounds that may be used for the
 CC treatment of S. epidermidis infections, e.g. endocarditis. AAH53971 to
 CC AAH5090 represent specifically claimed S. epidermidis genomic DNA
 CC AAH5098 represent oligonucleotide sequences and primers which are used
 CC in the exemplification of the present invention. AAH5091 to
 CC N.B. The present invention specifically claims all the polynucleotide
 CC sequences given in the sequence listing of the present specification,
 CC however the sequence listing only goes up to SEQ ID NO:4454 so even
 CC though sequences are given in the disclosure for SEQ ID NO:4465 to 4472,
 CC no sequences are present for SEQ ID NO:4455 to 4464.
 CC
 CC Sequence 910 AA:

Query Match 4.9%; Score 134; DB 22; Length 910;
 Best Local Similarity 24.6%; Pred. No. 0.019;
 Matches 98; Conservative 45; Mismatches 147; Indels 108; Gaps 23;

26 TPVTDELLANPPAGEMISYQONENRHSPLQIT-----ENV-----GOLQVWARQMO-76
 231 tptt-----knpltekgvgepektekvkqvdeivhyggeelkpgkdeifpnapsqg 286
 77 --PKR-----VOVPLIHGVMYLANPGDVIADAKGDLIMHRQL-PNIA--122
 287 dvppkpgvknptdtegvvtpvddvltkypvgdpl-----tsteepfdkkrefdpn 342
 123 --TINSGEPTRGMAIVANGVIVAGSTCOYSPFGCSGHSATGELMWRNYFIRAGEE 180
 343 tekvvkgep-----gtktitp-----tknpltekgv-----ge 373
 181 GDETWGNDYFARMGAMQITVDPYTNLHYGSTAVGPA--SETORGPGTLYGTNTR 238
 374 gept-----ekvtkqvdeivhyggeelkpgkdeifpnapsqgedvppk 419
 239 FAVR-PDTGEIYWRHQTLPRDNMOECTFEMAYTN--VDVOPSTMEGL-----QSTNP 289
 420 pgvknptdtegvv-----tpvdd-----vtkygvdgdpitsteelpfdkkrefdp 465
 290 NAATG-ERRVLTGP-CKT-GTMMQFADATGEFLMARDNNGMIESIDENGIVTVNEDA 346
 466 nlapgtekvkqkgpgetkltitpknpltekgvgepekteitkqvde--lvhygee 523

OY 347 ILKELDEVDVCPFLGGRDWS--AALNPDGITYFIP 382
 DB 524 lrpghkdefdpnapkgsgedvpgkpyknpdtgvevtr 561

RESULT 15
 AAG82914
 ID AAG82914 standard; Protein: 696 AA.

AC AAG82914;

DT 03-SEP-2001 (first entry)

DE S. epidermidis open reading frame protein sequence SEQ ID NO:2922.

XX Staphylococcus epidermidis SRI strain; infection; diagnosis;
 KM vaccination; endocarditis.

XX Staphylococcus epidermidis.

XX MO200134809-A2.

XX 17-MAY-2001.

XX 09-NOV-2000; 2000MO-US30782.

XX 09-NOV-1999; 99US-0164258.

XX (GLAX) GLAXO GROUP LTD.

XX Kimerly MJ;

XX WPI; 2001-316495/33.

XX DR N-PSDB; AAH53764.

XX Nucleic acids encoding polypeptides from Staphylococcus epidermidis,
 PT useful for vaccinating against infections, e.g. endocarditis -

XX Claim 18; Page 763-764; 2188pp; English.

XX AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides
 CC (II), given in AAG81454 to AAG83120, from Staphylococcus epidermidis.
 CC (I) and (II) can have antibacterial activity and therefore can be used
 CC in vaccination. The nucleic acids (I) may be used to produce the
 CC S. epidermidis polypeptides (II) via the production of vectors
 CC containing them which are used to produce hosts cells which express the
 CC polypeptides. The polypeptides (II) (and/or nucleic acids) may then be
 CC used to vaccinate subjects and to raise antibodies against the bacteria.
 CC The polypeptides may also be used to assay for other inhibitors of their
 CC activity and therefore identify compounds that may be used for the
 CC treatment of S. epidermidis infections, e.g. endocarditis. AAH53971 to
 CC AAH55090 represent specifically claimed S. epidermidis genomic DNA
 CC polynucleotide sequences from the present invention. AAH55091 to
 CC AAH55098 represent oligonucleotide sequences and primers which are used
 CC in the exemplification of the present invention.
 CC N.B. The present invention specifically claims all the polynucleotide
 CC sequences given in the sequence listing of the present specification,
 CC however the sequence listing only goes up to SEQ ID NO:4454 so even
 CC though sequences are present in the disclosure for SEQ ID NO:4465 to 4472,
 CC no sequences are present for SEQ ID NO:4455 to 4464.

XX Sequence 696 AA;

Query Match 4.9%; Score 133.5; DB 22; Length 696;
 Best Local Similarity 23.6%; Pred. No. 0.014;
 Matches 107; Conservative 56; Mismatches 148; Indels 143; Gaps 27;

OY 26 TPVTELLANPPAGEISTGQNOENTYRHSPLTQIT---ENV-----GQLDLVARGMO- 76
 DB 103 tptt---knplgtkevgsepteiktqpydeltteygeelkpgkhdefpnapksqge 158
 OY 77 --PGK-----VOVPLIHLDGVWYLANPGDVIGQIDAKTGDLIMEHRRQL-PNIA-- 122

DB 159 dvpqkpyvknptdtegvtrtppvdvltkygpydgdpj-----tsteipfdkkrtenplapq 214
 OY 123 --TLNSPEEPTRGMAIYANGVIYAGSTCOXSPFCFVSGHDSATGELMWNYPPIPAGEE 180
 DB 215 tekvvqkgep-----gkctltpt-----tknplgtgekv-----ge 245
 OY 181 GDETWGNDYEARMWTGAMGQITPDYPTNLVHYGSTAVGPA--SETORGITGGFLXGNTR 238
 DB 246 gept-----ekltkqpvdeivhyggeelkpgkhdefdpnapksgedvpgk 291
 OY 239 FAVR-PDTGEIYWRHQTLPDNDMDQECTFEEMATN---VDVQSTMEGL-----QSINP 289
 DB 292 pgvknptdtegvtr-----lppvdq-----vltkypvdqdsiltsteelpfdkkrdefcp 337
 OY 290 NATG-ERRVLTVGP-CKT-GTMOFDAETGEFLWARDJYONMIESIDENGIYTVNEDA 346
 DB 338 nlapgtkekvqgepgtkltltpttknplgtgekvgegepteiktqpyde--lvhygse 395
 OY 347 ILKELDEVDVCPFLGGRDWSAAL-----NPDG-IFYFIPLNVCYDMA 392
 DB 396 lrpghkdeid-----pnapvdsktevpgkpyknpdtgvevtrppvdvltkygk 444
 OY 393 VDOEFTSMD-----VYNTSNVTKLPFGKDMI 418
 DB 445 vgnplsteeipfdkkrvfnpd-----lkrgeerv 474

Search completed: May 24, 2002, 10:19:22
 Job time: 385 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2002, 10:20:13 ; Search time 39.53 seconds
(without alignments)
316.365 Million cell updates/sec

Title: WALICK-934-135.PEP
Perfect score: 2742
Sequence: 1 MKPTSLIMNSAGALALLAAP.....GMQVYVAGGVSYGSLINS 512

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: Issued Patents-AA:
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep.*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep.*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep.*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep.*
5: /cgn2_6/ptodata/2/1aa/PCTUS.COMB.pep.*
6: /cgn2_6/ptodata/2/1aa/Backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	484	17.7	754	US-09-296-284-4	Sequence 4, Appli
2	483.5	17.6	738	US-07-985-458-3	Sequence 3, Appli
3	475	17.3	720	US-09-296-284-25	Sequence 25, Appli
4	320.5	11.7	739	US-09-136-251-2	Sequence 2, Appli
5	116	4.2	3567	US-07-642-734C-4	Sequence 4, Appli
6	116	4.2	3567	US-08-439-009A-4	Sequence 4, Appli
7	113	4.1	1042	US-08-928-361B-5	Sequence 11, Appli
8	113	4.1	1837	US-08-928-361B-11	Sequence 5, Appli
9	107.5	3.9	1721	US-08-700-651-5	Sequence 10, Appli
10	105	3.8	1012	US-08-219-262B-10	Sequence 10, Appli
11	105	3.8	1012	US-09-031-655-10	Sequence 5, Appli
12	104.5	3.7	635	US-08-931-608A-5	Sequence 30, Appli
13	101.5	3.7	1043	US-08-928-361B-30	Sequence 6, Appli
14	101.5	3.7	1721	US-08-928-361B-6	Sequence 4, Appli
15	98.5	3.6	816	US-07-731-157A-4	Sequence 4, Appli
16	98.5	3.6	816	US-08-229-444B-2	Sequence 4, Appli
17	98.5	3.6	816	US-08-541-780-4	Sequence 13, Appli
18	97	3.5	574	US-09-142-623-13	Sequence 13, Appli
19	95.5	3.5	500	US-08-987-519-2	Sequence 2, Appli
20	95.5	3.5	548	US-08-931-605A-19	Sequence 19, Appli
21	95	3.5	512	US-08-216-276A-12	Sequence 2, Appli
22	94.5	3.4	824	US-09-356-818A-2	Sequence 3, Appli
23	94.5	3.4	824	US-09-626-589-3	Sequence 2, Appli
24	94.5	3.4	1612	US-08-169-927-2	Sequence 2, Appli
25	94	3.4	974	US-08-868-786-6	Sequence 6, Appli
26	94	3.4	983	US-09-394-200-2	Sequence 2, Appli
27	93.5	3.4	548	US-08-247-902A-2	Sequence 2, Appli

28	93.5	3.4	966	2	US-08-868-786-2	Sequence 2, Appli
29	93	3.4	380	3	US-08-971-782-4	Sequence 4, Appli
30	93	3.4	380	4	US-09-309-026-4	Sequence 2, Appli
31	93	3.4	459	3	US-08-971-782-2	Sequence 2, Appli
32	93	3.4	459	4	US-09-309-026-2	Sequence 2, Appli
33	93	3.4	659	4	US-09-626-589-2	Sequence 1, Appli
34	93	3.4	717	4	US-09-626-589-1	Sequence 6, Appli
35	92	3.4	532	2	US-08-883-534-3	Sequence 3, Appli
36	92	3.4	606	2	US-09-204-764-3	Sequence 2, Appli
37	92	3.4	606	3	US-08-038-682-2	Sequence 2, Appli
38	91.5	3.3	1536	1	US-08-302-832-2	Sequence 2, Appli
39	91.5	3.3	1536	2	US-08-530-198-2	Sequence 2, Appli
40	91.5	3.3	1536	2	US-08-469-850-2	Sequence 2, Appli
41	91.5	3.3	1536	2	US-08-728-470-2	Sequence 2, Appli
42	91.5	3.3	1536	2	US-08-617-697-2	Sequence 2, Appli
43	91.5	3.3	1536	4	US-08-617-697-2	Sequence 2, Appli
44	91.5	3.3	1536	2	US-08-824-707-2	Sequence 2, Appli
45	91	3.3	306	2	US-08-824-707-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-09-296-284-4
; Sequence 4, Application US/09296284A
; Patent No. 6204040
; GENERAL INFORMATION:
; APPLICANT: Chol, Eun-Sung
; APPLICANT: Rhee, Sang-Ki
; APPLICANT: Lee, Eun-Hae
; TITLE OF INVENTION: Glucosyltransferase, Genes
; TITLE OF INVENTION: And Methods of Use Thereof
; FILE REFERENCE: 1533.08/0000
; CURRENT APPLICATION NUMBER: US/09/296.284A
; CURRENT FILING DATE: 1999-04-22
; NUMBER OF SEQ ID NOS: 87
; SOFTWARE: Patent Ver. 2.0
; SEQ ID NO 4
; LENGTH: 754
; TYPE: PRT
; ORGANISM: Gluconobacter suboxydans
; US-09-296-284-4

Query Match	17.7%	Score 484	DB 4	Length 754
Best Local Similarity	25.9%	Pred. No. 2.9e-36		
Matches 156	Conservative	81	Mismatches 239	Indels 126
Gaps 16				
14	LALIAAPAFQAQVTPV-----TDELLANPPAGEWISGQNDENRRHSPLQITTEWVGQ	66		
17	LGCAALAFATSPVALAEDTGATLTNADQHPGDMWSYSEORTSPDQITKDNASN	76		
67	LQLVWANGMPKGVQV-TPLIHGVMVYLANPDGVICADIAIDAKTGLIWEHRQRP-NIATL	124		
77	LKLAHMYDLPTNRGQBSTPLVDGVWATTTNMSKKALDAATGKILMSYDPRVGINADR	136		
125	NSFGEPTRGMAI-----	136		
137	GCCTTVKRGAAVWNGKYFFGTEDGRILALDAKTGLKLSVTVPEADLGHORSTYDGA	196		
137	--VANGVIVAGSTQYSPFGC--FVSGHSATGEELMRVYFIPRAGEEGD-----	182		
197	PRIAKGVIILN--GGAFFGARGVTAIDETKMKMRFFYTPNDNKPDAASDVILMS	254		
183	---ETWGDYDARMWTC---AMGQITVDPVNTLVHYGSTAVGPASTQGTGCGTLTGN	236		
225	KAYPTWGGGAMKQGGGGTWDLSIYDPTDLVYLGVGNSFWNKKFSEKGNLFLG	314		
237	TRFAVRPDTGETIWRHQTLPDNDNDQCTEFEMMTNVYDQPTSTEMGLQINPNATGE-	295		
315	SIVAINPDTGKYVWHFDETPMDQMDYTSVQOIHALLDMPV-----NGEM	357		

FEATURE:

453 LKNGTDRYFRAISQETGETLMQTRLATVASSQAISYEVDGMQYVAIAGG 502

```

RESULT 3
US-09-296-284-25
; Sequence 25, Application US/09296284A
; Patent No. 6204040
; GENERAL INFORMATION:
; APPLICANT: Choi, Eun-Sung
; APPLICANT: Rhee, Sang-Ki
; APPLICANT: Lee, Eun-Hae
; TITLE OF INVENTION: Glucanobacter Suboxydans Sorbitol Dehydrogenase, Genes
; TITLE OF INVENTION: and Methods of Use Thereof
; FILE REFERENCE: 1533 0870000
; CURRENT APPLICATION NUMBER: US/09/296,284A
; NUMBER OF SEQ ID NOS: 87
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 25
; LENGTH: 720
; TYPE: PRT
; ORGANISM: Glucanobacter suboxydans
US-09-296-284-25

```

```

Query Match 17.3%; Score 475; DB 4; Length 720;
Best Local Similarity 25.7%; Pred. No. 1.8e-35;
Matches 150; Conservative 82; Mismatches 229; Indels 122; Gaps 16;

OY 26 TPVDELLANPPAGEWTSYGNQENYRHSPLTQITTEVGOLOLVYARQMOQYV-TP 84
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
   5 TAITN---ADHPDQMSWSTGRTYSEQKXSPLDQITKDNMSNLKLAHYDLDTNRQEGTP 61
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
   OY 85 LHHGVWYLANPGVYIAQIDAKTGLIWEHRRLP-NIALNFGSEPTGML----- 136
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
   62 LVDGVWATWMSKMLDAATGKILMSYDPKVFNIADRGCCDVTNRCAVWNSKYVF 121
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
   OY 137 -----VANGVIAGSTCOYSPF 153
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
   DB 122 GFEDGLALDAKGLKLYWSYTYPKKQLOHQRSYTVDAQPRIAKGVILGN--GGAEF 179
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
   OY 154 GC--FVSGHDSATGELMKNRYFIPRAGEGD-----ETGNDYEAHMMWTG-- 196
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
   DB 180 GAAGFTATADETKGMDRFFTVPNPDKPDGAASDDVLSKAYPTWGGKCAKQGGGG 239
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
   OY 197 -ANGQITTPYTNLVHYGSAVGPASETQGRPGGLTYGNTREAVRPDGTGLVWRHQT 255
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
   DB 240 TVWDSLIYPVDLYLVGVGNGSPWNYKFRSGKGNLFLGSYVALINDPCKYWHFOET 299
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
   OY 256 PRDMWDECFEFEMATVNVQVSTEMEGLQSIINPAATGE-RVYLGVPCKTGMQFPA 314
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
   DB 300 PMQMDITTSVQIMALDMYV-----NGEMRHVLYNAP-KNGFFYIIDA 341
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
   OY 315 ETGELLARDTNYQNNIESIDE-NGIIVYNEDAILKELDVEYDVCPTFLGGRDPSAALN 373
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
   DB 342 KTGKFIISGRKYTYENMANGLDPTVGRPNYDMLTKNCKPMTYGIPLDGGHNFAAAYS 401
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
   OY 374 PDGSLYFPLNNVCY-----DMAVDOETSMYNTSN---VTLPKPKD 416
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
   DB 402 PQTLLVYIPAOQVYFVDPKQGFKAHDSWNLGLDMNKIGLIDNDPHKADKQFLKD 461
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
   OY 417 MIGRIDALIDISTGHTLMSVERAANSPVLSSTGGVLFNGGDRYFRALSQETGELMOT 476
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
   DB 462 LKGWIVADWPQKQAAFTVYHKGKPMWNGGLATAGVLFQGLANGEPHAYDATITGDTTF 521
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
   OY 477 RLATVASQAISYEVDGMQYVAI-----AG-----GGVSYGSG 509
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
   DB 522 PAQSAIIPAYTYTANGKQYVAIVEGVNGGIYPFELGCVARTSG 564
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1

```

```

APPLICANT: HOSHINO, Tatsuo
APPLICANT: MIYAZAKI, Taro
APPLICANT: OHTA, Setsuko
APPLICANT: SHINJOH, Masako
APPLICANT: TOMIYAMA, No. 61271561buni
TITLE OF INVENTION: D-Sorbitol Dehydrogenase Gene
FILE REFERENCE: D-Sorbitol Dehydrogenase Gene
CURRENT APPLICATION NUMBER: US/09/136,251A
CURRENT FILING DATE: 1998-08-19
EARLIER APPLICATION NUMBER: EP 97114432.4
EARLIER FILING DATE: 1997-08-21
NUMBER OF SEQ ID NOS: 9
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 2
LENGTH: 739
TYPE: PRT
ORGANISM: Glucanobacter suboxydans
FEATURE:
NAME/KEY: SIGNAL
LOCATION: (1)...(24)
US-09-136-251-2

```

```

Query Match 11.7%; Score 320.5; DB 3; Length 739;
Best Local Similarity 21.6%; Pred. No. 4.3e-21;
Matches 150; Conservative 79; Mismatches 186; Indels 281; Gaps 33;

OY 23 AQYTPVDELLANPPAGEWTSYGNQENYRHSPLTQITTEVGOLOLVYARQMOQYV-TP 78
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
   86 SQVPYAPAPQASANPARGWAVAYGRDDHQTRYSPLSEITPENASKLKAIFYHTGSPRG 145
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
   OY 79 KV-----QVTPV-HDGVWYLANPGDVIAQIDAKTGLIWE----- 113
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
   DB 146 QVKKMAAETPLVIGDGLTCSAMNDIIR-LDPATGKQIMRNVYKHSIPTACKGV 204
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
   OY 114 -----HRRQL-----PN-----IATLNSFGEPTGM 134
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
   DB 205 TYFTSSVPEGQCHNRRLIEGLDMRLIANDAEFGDFPCNFGHGQVNLMLGLESYPGF 264
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
   OY 135 AI-----VANGVIAGSTC-----QYSPGCFVSGHDSATGELMKNRYFIPRAGEGD 183
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
   DB 265 VSMTPAPPVINGVYVNVNHEVLDGRRMAPSG-VIRGYDAESGKFW----- 309
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
   OY 184 TWGNDYARWMTGAMQOITVDPYTNLVHYGSAVGPASETQGRPGGLTYGNTREAVRPDGTGLVWRHQT 238
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
   DB 310 AMDVNNSGR-----SOPAY-RVTVITAVERRILGILPDRRQK-----GRLRPDNS 355
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
   OY 239 -----FAVRPD-----TGLVWRHQTLPDMNDQDCEFMATVNVQVPS 278
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
   DB 356 AADYYSALRSDAENKVSAAVAIDVKTGSPRVYFQTAHNDVMDYDIGSATL----- 407
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
   OY 279 TEMEGLOSINPNAATGERVLTGVPCKTGMQFPAETGELLARDTNYQNNIESIDE- 337
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
   DB 408 MDMPG-----FDGQIVPALIM--PTRKQTFVLDRTCK-----PILVEERF 448
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
   OY 338 ---GIV-----TVNDAI-----LKELD----- 352
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
   DB 449 APSGVIPQDPSPTQPSVGMALRYVDLKEIDMGMKSIDQLCFKIFRRANYGEPT 508
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
   OY 353 -----VEYDVCPTFLGGRDPSAALNDPSGIYFPLN-NCYDMAVDOETSM--- 400
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
   DB 509 PPSVDKFWLEY--PGYNGSGDMSMSYDQSGILIANWNTPEYDOLVYRKRAADSLGLM 565
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
   OY 401 -----DVYNTSNVTKLPKCKDMIGRIDALIDIS 427
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
   DB 566 PIDDPNFKPGGGAGBANGAMDGPYGIYVTPWQY-TGMCKNRP---YGMITAIIDMK 620
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
   OY 428 TG-RTLMS-----VERAANSPVLSSTGGVLFNG-CTDRYFRALS 466
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
   DB 621 HGOQVLMQHPGLGTARANGPWLPTGLPWEIGTPNNNGSVVGGILFTGAATDMQRAID 660
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
   OY 467 QETGELMOTRLATVASQAISYEVDGMQYVAIAGG 502
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1

```

```

RESULT 4
US-09-136-251-2
; Sequence 2, Application US/09136251A
; Patent No. 6127156
; GENERAL INFORMATION:

```

Db 681 EHTGKVVSAVLPGGGQANPMTEANCHOYVALAGG 716

RESULT 5

US-07-642-734C-4

Sequence 4, Application US/07642734C

Patent No. 5824513

GENERAL INFORMATION:

APPLICANT: Katz, L

APPLICANT: Donadio, S

APPLICANT: Mcalpine, J B

TITLE OF INVENTION: Recombinant DNA Method for Producing

TITLE OF INVENTION: Erythromycin Analogs

NUMBER OF SEQUENCES: 27

CORRESPONDENCE ADDRESS:

ADDRESSEE: Edward H. Gorman

STREET: Abbott Laboratories D377/Abpd-2 One Abbott

CITY: Abbott Park

STATE: IL

COUNTRY: US

ZIP: 60064-3500

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/642,734C

FILING DATE: 17-JAN-91

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Daughters, Andreas M

REGISTRATION NUMBER: 32652

REFERENCE/DOCKET NUMBER: 4952.US.01

TELECOMMUNICATION INFORMATION:

TELEPHONE: 708-937-9396

TELEFAX: 708-938-2623

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 3567 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-07-642-734C-4

Query Match 4.2%; Score 116; DB 2; Length 3567;

Best Local Similarity 23.4%; Pred. No. 0.55;

Matches 127; Conservative 47; Mismatches 232; Indels 136; Gaps 25;

21 ARAQVPTVDELINPAGEW-----ISYQONQENRHSPLTQITTEVNGQLQYV 71

2911 AFAPIAVTDHRLARVPDG-MSDADAAPVIAV-----TTAHYALHDLAAGSGLIH 2964

72 ARGMPGKVVQVPLIHGVMYLANPGDVIAIDAKTGLDW-----EHRRLPNATLNS 126

2965 AAAGVGMAAVALARRAGAEVLAATAG-----PAKHGTLRALGLDEHIIASSRETGFARK 3018

127 FGEPT--RGAIVAN--GVIVASTQYSPFGCFVS-GHDSATGELMKNFIP-RAGE 179

3019 FRERTGGRGVVPLNSLIGELDLDSADLLAEDVFEVEMGKTDLRDAGDFRGYAPFDLGE 3078

180 EGDTEWMDY-EARWMTGANGQITTYDFVNLV-----HYGSAFV-GRASE 222

3079 AGDRDLGELIREVVGLGA-GEIDLPLFVSAMWELGSAPALQMSGRHVGKLVLPQAPV 3137

223 TQGT--PGTLYGTTRFAVRPDGELVYWRHQL-PRDNDQDCTEEMAVTVNDVQPS 278

3138 DPDCVTLTGT--GTIGRLAHVTEHGVRLLLVSRGADAPGDELAEIEDIGAS 3195

279 TEMGLOSINPNAT-----GERRVLNGVCKTGMQFDEATGETELMARDTNYQMIESI 334

Db 3196 AETACDTRADNALSALDGLPRELTGVVHAAGVL----- 3330

QY 335 DENGIVT-VNEDAILKELDYEDVCPFFLGRRMPSAALNPDGIFITPLNNVCYDMMAV 393

Db 3231 -ADGLVTSIDEPVIEQVLRKAVDA-----WNLHELTANTGJSFFVLFSASAVLA- 3280

QY 394 DQFTSMVYITSNV-----TKLPBGK-----DMGRIDAID 425

Db 3281 ---GGQGVYAANESINLAALRRTRGIPAKALGWLMAQASMTSGLDRIAR----- 3332

QY 426 ISTGRTLSVRAAANSPVLTSGGVPLNGGTDRTFRALSOETEFTL-WQRLATVASC 484

Db 3333 --TGVAALPTEPRALALEFDSLRRGGEVVPPLINSALRAAEFVEVLRGVRAKLRBAAG 3390

QY 485 QA 486

Db 3391 QA 3392

RESULT 6

US-08-439-009A-4

Sequence 4, Application US/08439009A

Patent No. 6004787

GENERAL INFORMATION:

APPLICANT: Donadio, S

APPLICANT: Katz, L

APPLICANT: Mcalpine, J B

TITLE OF INVENTION: Method of Directing Biosynthesis of

TITLE OF INVENTION: Specific Polypeptides

NUMBER OF SEQUENCES: 27

CORRESPONDENCE ADDRESS:

ADDRESSEE: Steven F. Weinstein

STREET: Abbott Laboratories D377/Abpd-2 One Abbott

CITY: Abbott Park

STATE: IL

COUNTRY: US

ZIP: 60064-3500

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/439,009A

FILING DATE: 11-MAY-1995

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Casuto, Dianne

REGISTRATION NUMBER: 40,943

REFERENCE/DOCKET NUMBER: 4952.US.D1

TELECOMMUNICATION INFORMATION:

TELEPHONE: 847-938-3137

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 3567 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-439-009A-4

Query Match 4.2%; Score 116; DB 3; Length 3567;

Best Local Similarity 23.4%; Pred. No. 0.55;

Matches 127; Conservative 47; Mismatches 232; Indels 136; Gaps 25;

21 ARAQVPTVDELINPAGEW-----ISYQONQENRHSPLTQITTEVNGQLQYV 71

2911 AFAPIAVTDHRLARVPDG-MSDADAAPVIAV-----TTAHYALHDLAAGSGLIH 2964

72 ARGMPGKVVQVPLIHGVMYLANPGDVIAIDAKTGLDW-----EHRRLPNATLNS 126

2965 AAAGVGMAAVALARRAGAEVLAATAG-----PAKHGTLRALGLDEHIIASSRETGFARK 3018


```

1      LENGTH: 1042 amino acids
2      TYPE: amino acid
3      STRANDEDNESS:
4      TOPOLOGY: linear
5      MOLECULE TYPE: protein
6      US-08-928-361B-11
7
8      Query Match          4.1%; Score 113; DB 3; Length 1042;
9      Best Local Similarity 21.1%; Pred. No. 0.14;
10     Matches 139; Conservative 75; Mismatches 224; Indels 222; Gaps 36;
11
12     25 VTPTVDLLANPPAGEMISTYQNOE-----NYHSLPQTITTEVNGQGLVWARGMOP 77
13     : : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
14     Db IDPTTG-LFPNPPPGCHLINPTNNNTDSSFAQAYKRAVSNGIKIDNVYGLPVDITLPR 359
15
16     78 GKQVTPLIHDGVWYLYANPGDVIOAIDAKTGLIMEH-----RROLPIATLN-SRFG 129
17     : : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
18     Db DPVSDIP-----FNSITGEL--VDPSTGKPINNYAGIVSGKRGLPIEDENGINLFD 409
19
20     130 PTRGMAIVANGVI-----VAGST-----CQISPECFPSGSDASTGELMRANFIPRAG 178
21     : : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
22     Db PSTKPLIDGNNOALVNPETNSTVSGSITSGTKRP-GIPVNGGVYPDEE-----AK 459
23
24     179 EEGDPTNODXEAMRMGANGQILYDPVTILVYGSTA--VGPASETORGTG-----229
25     : : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
26     Db DQADK--GKD--GLIYPPNTSINKDPVYTOYSNNTGNINP--ETGKVLPGSLPGSLN 512
27
28     230 -----GTLGNTFRPAVRPTDGEIYVHRKQTLPRDN--WDOECTFPMWYTN 272
29     : : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
30     Db YPSFMTPOQTEIIRGKPDVITVGLPYDSTGELIDPATKLPISVAGDELLLEVANIIT 572
31
32     273 -----VDVO-----PSTMEGLDS-----INP-----NAATG-----ERRV 298
33     : : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
34     Db DEVTLPIDLETLGELPRDPVSGLPOLPNCGLVDPSNNKKPIPGSHSGFINGTSGEOSHERDP 632
35
36     299 LNVGPKCTGTMMQFD-----AETGEFLMARDPYTNQNNI--ESIDENGIYTVENDAIL 348
37     : : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
38     Db SGKGLPDPMTGLPFDEDSGLINPETGDKLQSGHSFTMPVPGKPOGENGIMTPEO-IL 691
39
40     349 KELDV-----EVDVCP-----TFL--GGRDWPSAAL-----372
41     : : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
42     Db EALNKLPTISNEVNISSPRSSDAVPRPRTNWMNKSIGCTQVGVGKTIIPSAASVITHAL 751
43
44     373 -----NPDGIFYFILNANCYDMAAVDQ-----ETSADYYNINNTKLP 413
45     : : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
46     Db GPTQTDPTTGLPSPSTGIPFIPGFNVLVDPQGEQIKGSVPVSLIYVENKIVTEAAY 811
47
48     414 GKRMIGRIDAIDISGRITLMSYERAAAYSPVLSNG-----GGVLENGIDTRYFA--464
49     : : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
50     Db 812 GL-----PVDPKTEGFIPI-----SYLPFAKNGELIDPISGKYSSGIAGFISGKAG 859
51
52     465 -----LSQETGETLMQF-----RLATVASGOAISYEYDGMQYIALAG-----GGV 504
53     : : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
54     Db SQSKSDSESCNPIDPSTNMPYDKTGKILIDPESGKIAIDNSVSGV-FATVPGTAAPKKG 918
55
56 RESULT      8
57 US-08-928-361B-5
58 : Sequence 5, Application US/08928361B
59 : Patent No. 6071518
60 :
61 : GENERAL INFORMATION:
62 : APPLICANT: Petersen, Carolyn
63 : TITLE OF INVENTION: PEPTIDES, POLYPEPTIDES, GLYCOPROTEINS,
64 : TITLE OF INVENTION: THEIR FUNCTIONAL MUTANTS, VARIANTS, ANALOGS AND FRAGMENTSS
65 : TITLE OF INVENTION: FOR TREATMENT AND DETECTION/DIAGNOSIS OF CRYPTOSPORIDIUM
66 : TITLE OF INVENTION: SPECIES INFECTIONS
67 : NUMBER OF SEQUENCES: 30
68 : CORRESPONDENCE ADDRESSES:
69 : ADDRESSEE: PETERS, CERNY, JONES & BIKSA
70 : STREET: 385 Sherman Avenue, Suite 6
71 : CITY: Palo Alto
72 : STATE: CA

```

COUNTRY: USA
 ZIP: 94306-1840
 COMPUTER READABLE FORM:
 MEDIUM TYPE: floppy disk
 OPERATING SYSTEM: IBM PC compatible
 SOFTWARE: Patent Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/928,361B
 FILING DATE: 12-SEP-1997
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 60/026,062
 FILING DATE: 13-SEP-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: VERNY, HANA
 REGISTRATION NUMBER: 30,518
 REFERENCE/DOCKET NUMBER: 480.76-1(HV)
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 650-324-1677
 TELEFAX: 650-324-1678
 INFORMATION FOR SEQ ID NO: 5:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1837 amino acids
 TYPE: amino acid
 STRANDEDNESS:
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-928-361B-5

Query Match 4.1%; Score 113; DB 3; Length 1837;
 Best Local Similarity 21.1%; Pred. No. 0.36; Mismatches 224; Indels 222; Gaps 36;
 Matches 139; Conservative 75;

25 VPPVDELAPNAPAGWISYGNQ-----NYRHSPLQITTEVNGOQLVWARGMOP 77
 1096 IDPTTG-LPFPNPTGHLINPTNNNTMDSFAGAYKIVASNGIKTDVNGLPVDITGLPK 1154
 78 GKQVTPPLIHGVMYLANPGDVI-----QADAKTGLD-----IWEHRRLPNIAF 123
 1155 DEVSDIP-----FNSTTGL-----VDPSTGKPINNTAGIVSGKRGRLPIEDENGMLFD 1204
 130 PPRGAIVANGVI-----VAGST-----COYSPGCFVSGHDSATGEELMRNFIP-- 175
 1205 PSTKLPIGNNQVLNPETNSVSGSTSGTKPKR-GIPVNGGVVPEE-----AK 1254
 179 BEGGDETMGNDYEAARMGTAMGQIYDPTNLVHYGSTA--VGPASETORGTPG----- 229
 1255 DOADK--GKD--GLIYPTNSINKDPVTNTQYSNNTGNIINP--ETGKVIYPSLPSLS 1307
 230 -----GLYGTNTRFAVAPDGTGEIYVWRHQTLPKRN--WDQECTFEKMA 269
 1308 YPSNTPQQTDELITGKRPVDVTGLPYDPSTGEIIDPATKLPISGVAAGDEILLEVLI 1367
 273 -----VDVO-----PSTEMEGLOS-----INP-----NATG-----ERRV 298
 1368 DEVTGLPIDLETGLPRPVSGLPQLPENGILVDSNKKPIPGSHSGFINSTSGSHEKRP 1427
 299 LTVGPKCTGMWOFD-----AETGEFLMADNTYQNNI--ESTIDENGIVYVNDATL 348
 1428 STGKPLDPNPGGLPDEDSGLINPETGDKLQSGHSGTFMPVPGKPOGEGGIMTPEO--TL 1486
 349 KELDY-----EYDVC-----FEL--GGGDMPSAL----- 372
 1487 EALNKLPTSNENVISPPSSDAVDPRTNTMWNKISGQYQVDGKKTIPGSAASYIHIAL 1546
 373 -----NDPSGIYFIPLANNVCYDMAAVDQ-----EFTSMDYNTSNVTKLPP 413
 1547 GTPTQVDDPTTGLPSDPTGFLPIFGFNVVLVDPQTGEQIGSVYVSLVYKEKNITTEAY 1606
 414 GKDMIGRIDAIIDISTGRITMSVRAAANYSPVLSG-----GVLFPNGGIDRIFRA--- 464

Db 1607 GL-----PYDPKTEGPIDPI-----SYLPAKNGELIDIPISGKYESGSIANGFISGKAG 1654
 QY 465 -----LSQFTGETLMQT-----RLATVASGQAISEVDSQMOVALAG-----GGV 504
 Db 1655 SQSKSDSDSGNPIDPSTNMFPYDKRTGKLIDPESGAIIDNSVSGV-FATVPGTAARPKGGV 1713
 RESULT 9
 US-08-700-651-5
 ; Sequence 5, Application US/08700651B
 ; Patent No. 6015882
 ; GENERAL INFORMATION:
 ; APPLICANT: PETERSEN, CAROLYN
 ; APPLICANT: LEECH, JAMES
 ; APPLICANT: NELSON, RICHARD, C.
 ; APPLICANT: GUY, JIRI
 ; TITLE OF INVENTION: VACCINES, ANTIBODIES, PROTEINS, GLYCOPROTEINS, DNAS AND RNAS
 ; TITLE OF INVENTION: FOR PROPHYLAXIS AND TREATMENT OF Cryptosporidium parvum
 ; FILE REFERENCE: 480.19-4(HV)
 ; CURRENT APPLICATION NUMBER: US/08/700,651B
 ; CURRENT FILING DATE: 1997-08-14
 ; EARLIER APPLICATION NUMBER: 08/415,751
 ; EARLIER FILING DATE: 1995-04-03
 ; NUMBER OF SEQ ID NOS: 15
 ; SOFTWARE: Patentln Ver. 2.0
 ; SEQ ID NO 5
 ; LENGTH: 1721
 ; TYPE: PRT
 ; ORGANISM: Cryptosporidium parvum
 US-08-700-651-5

Query Match 3.9%; Score 107.5; DB 3; Length 1721;
 Best Local Similarity 20.6%; Pred. No. 1.1;
 Matches 137; Conservative 75; Mismatches 222; Indels 231; Gaps 37;

25 VPPVDELAPNAPAGWISYGNQ-----NYRHSPLQITTEVNGOQLVWARGMOP 77
 979 IDPTTG-LPFPNPTGHLINPTNNNTMDSFAGAYKIVASNGIKTDV-----YGLPV 1029
 78 GKQVTPPLIHGVMYLANPGDVI-----QADAKTGLD-----IWEHRRLPNIAF 123
 1030 GEITGLP-----KPGSDIPFNSTTGLVDPSTGKPINNSTAGIYSGKRGLEPID 1080
 124 IN-SPEPTFGMAIVANGVIYASSTCOYSPGCFVSGHDSAT-----GEELMRNFIP-- 175
 1081 ENGNLFPDSTMLPIDGNNQVLNPET-----NSTVSGSTSGTKPKRGPVNGGVVPEE 1134
 176 RAGEGDETMGNDYEAARMGTAMGQIYDPTNLVHYGSTA--VGPASETORGTPG----- 229
 1135 EAKDQADK--GKD--GLIYPTNSINKDPVTNTQYSNNTGNIINP--ETGKVIYPSLPS 1187
 230 -----GLYGTNTRFAVAPDGTGEIYVWRHQTLPKRN--WDQECTFEKMA 269
 1188 SLNTPSNTPQQTDELITGKRPVDVTGLPYDPSTGEIIDPATKLPISGVAAGDEILLEVLI 1247
 270 VTN-----VDVO-----PSTEMEGLOS-----INP-----NATG-----E 295
 1248 ITTDEVYGLPIDLETGLPRPVSGLPQLPENGILVDSNKKPIPGSHSGFINSTSGSHE 1307
 296 RVLITGVPCCTG--MWOFD-----AETGEFLMADNTY-----Q 328
 1308 KDPSTGKPLDPNPGGLPDEDSGLINPETGDKLQSGHSGTFMPVPGKPOGEGGIMTPEO-- 372
 329 NMIESIDENGIVYVNDATL-----DAI-----KLELDYEVYCTPTFAG----- 364
 1368 QILEAL--KRLPTSNENVISPPSSDAVDPRTNTMWNKISGQYQVDGKKTITIGSASV 1425
 365 -----RDWPSAL--NDPSGIYFIPLANNVCYDMAAVDQ-----EFTSMDYNTSNV 408
 1426 IHTALGTPTQVDDPTTGLPSDPTGFLPIFGFNVVLVDPQTGEQIGKSVYVSLVYKEKNIV 1485

QY 409 TKLPCKMDIGRIDAIDISTGRTILMSVERAANYSPVLSTG-----GGVLENGCIDRIF 462
 DB 1486 TEAAVGL-----PVDRKTGFPIDP-----SYLPEAKNGELIDPISGXFSGSINGFI 1533
 QY 463 RA-----LSQFTEGTLMTQ-----RLATVAGSQAISEYVGMQVVAIAG----- 501
 DB 1534 SGKAGSQSKSDSEGNIPDPTNMPDPRKGLIDPESGIAIDNSVSGV-FATVPGTAAAP 1592
 QY 502 --CGV 504
 DB 1593 KRGGV 1597

RESULT 10
 US-08-219-262B-10
 ; Sequence 10, Application US/08219262B
 ; Patent No. 5788970
 ; GENERAL INFORMATION:
 ; APPLICANT: VAKHARIA, VIKRAM
 ; APPLICANT: SNYDER, DAVID B
 ; APPLICANT: MENDEL-WHERSAT, STEPHANIE A
 ; TITLE OF INVENTION: CHIMERIC INFECTIOUS BURSAL DISEASE VIRUS
 ; TITLE OF INVENTION: CDNA CLONES, EXPRESSION PRODUCTS AND VACCINES BASED
 ; TITLE OF INVENTION: THEREON
 ; NUMBER OF SEQUENCES: 15
 ; CORRESPONDENCE ADDRESSES:
 ; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT
 ; STREET: 1755 S. JEFFERSON DAVID HIGHWAY, FOURTH FLOOR
 ; CITY: ARLINGTON
 ; STATE: VIRGINIA
 ; COUNTRY: USA
 ; ZIP: 22202
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: FLOPPY disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/219,262B
 ; FILING DATE: 29-MAR-1994
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: OBLON, NORMAN F
 ; REGISTRATION NUMBER: 24,618
 ; REFERENCE/DOCKET NUMBER: 2747-047-27
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (703) 413-3000
 ; TELEFAX: (703) 413-2220
 ; TELEX: 248855 OPAT UR
 ; INFORMATION FOR SEQ ID NO: 10:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1012 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS:
 ; TOPOLOGY: unknown
 ; MOLECULE TYPE: protein
 ; ORIGINAL SOURCE:
 ; ORGANISM: Infectious bursal disease virus
 ; STRAIN: OH
 ; US-08-219-262B-10

Query Match 3.8%; Score 105; DB 1; Length 1012;
 Best Local Similarity 22.2%; Pred. No. 0.77;
 Matches 84; Conservative 45; Mismatches 130; Indels 120; Gaps 19;

QY 206 VTNLVHYSSTAY-----GRAS-----ETGRTGPGGTLGTNTRF 239
 DB 1 MTNLMHDTQOIVFIRSLMPTGPASIPDPTLEKHTLSRSTYNTLVGDTGSLIVFF 60
 QY 240 AVRPDGEIYVWRHQTLPD-----NMDECTEFEMAVT-----NVDQPSFE 280
 DB 61 PGFP--GSVVGAIYTIQNSNSYOFD-----MLLTAONLPVSYNYCRILVSRSLIVRSSTL 113

QY 281 MEGLOSINP--NAATGERRVLGVPCKTGTMTMOPDAETGEFLNARDNTYONMTESI----- 334
 DB 114 PGGVALNCTINAVT-----FQGISLEL-----TDVSYNGLSMANTANIN 152
 QY 335 DENGIVTVNEDAILKELDVEIDVCTPFLGGDRMPSAALNPDSGIFYFLPUNVC----- 387
 DB 153 DKIGNVLVGEQVTVLSLPTSYDLSTYRLGD--PIPAAGIDP-----KLATCCSSDRPR 204
 QY 388 -YDMAAVDQ--EF-----TSMQVNTSNVTKLPFGKDMI-GRIDAIDISTGRTILM- 433
 DB 205 VYTVTRADYQFSSQLISGCVTTTLFTANIDALISLSVGGELIFSQYTIHSIEVDVYIIF 264
 QY 434 -----SVERAANYSPVLSTGGVLEFN--GGTDYFRALSDQETGTLMTQTLAVAS 483
 DB 265 IGPDGEYVYKAVADDFGLTGTNNILVPFNILGPTSEITOPTSMKLEVVYKKRGGT--A 322
 QY 484 GOAISTEVDGMQVVAIAG 502
 DB 323 GDPISWTVSGTLAVTVIGG 341

RESULT 11
 US-09-031-655-10
 ; Sequence 10, Application US/09031655
 ; Patent No. 6017759
 ; GENERAL INFORMATION:
 ; APPLICANT: VAKHARIA, VIKRAM
 ; APPLICANT: SNYDER, DAVID B
 ; APPLICANT: MENDEL-WHERSAT, STEPHANIE A
 ; TITLE OF INVENTION: CHIMERIC INFECTIOUS BURSAL DISEASE VIRUS
 ; TITLE OF INVENTION: CDNA CLONES, EXPRESSION PRODUCTS AND VACCINES BASED
 ; TITLE OF INVENTION: THEREON
 ; NUMBER OF SEQUENCES: 15
 ; CORRESPONDENCE ADDRESSES:
 ; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT
 ; STREET: 1755 S. JEFFERSON DAVID HIGHWAY, FOURTH FLOOR
 ; CITY: ARLINGTON
 ; STATE: VIRGINIA
 ; COUNTRY: USA
 ; ZIP: 22202
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: FLOPPY disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/031,655
 ; FILING DATE:
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/219,262
 ; FILING DATE: 29-MAR-1994
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: OBLON, NORMAN F
 ; REGISTRATION NUMBER: 24,618
 ; REFERENCE/DOCKET NUMBER: 2747-047-27
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (703) 413-3000
 ; TELEFAX: (703) 413-2220
 ; TELEX: 248855 OPAT UR
 ; INFORMATION FOR SEQ ID NO: 10:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1012 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS:
 ; TOPOLOGY: unknown
 ; MOLECULE TYPE: protein
 ; ORIGINAL SOURCE:
 ; ORGANISM: Infectious bursal disease virus
 ; STRAIN: OH
 ; US-09-031-655-10

Query Match 3.8%; Score 105; DB 3; Length 1012;
 Best Local Similarity 22.2%; Pred. No. 0.77;
 Matches 84; Conservative 45; Mismatches 130; Indels 120; Gaps 19;

QY 206 VTNIYHGSTAV-----GPAS-----ETGRTPGTLYGTNTRF 239
 Db 1 MTNLMDDHQIIVPIRSLMLPTTGPASIPDDTEKHTLRSETSTYMLTVDGDSGLIVPF 60
 QY 240 AVRDGTGEIWRHQTLPD---NMDOCTEEMAVT-----NVDQPSPE 280
 Db 61 PGFP--GSVGAHTTLOSNGSYQPDQ---MLTRQNLPSVYNGRLVRSRLTVRSSTL 113
 QY 281 MEGQSINP--NAATGERRLVGVPCKTGMOPDAETGEFLMARDTNYNNIESTI--- 334
 Db 114 PGVYALNGTINAVT-----FQSLSEL---IDSYNGLSMTATNIN 152
 QY 335 DENGIYVNDALIKELDEYDYCPTFLGRDMPASALMPDSGIYFIPLNVC----- 387
 Db 153 DKGNVLVGGVYVLSLPTSTLDLYRLD-PIPAAGLDP-----KIMATCDSDRPR 204
 QY 388 -YDMAVDO-EF-----TSMQVNTSNVTKLPKQMI-GRIDAIDISTRTLM- 433
 Db 205 VYVTVTADEXQFSSQLIPSGVKTTLTANIDALISLVGGELEFSQVTHISTEDVTLYF 264
 QY 434 -----SYERAANYSPVLSGGVLFN--GTDYFRFALSQETGELMOTRLATVAS 483
 Db 265 IGFDTGEVTVKAVADEGLTGTNNMLYFPNLGPGTSEITPITSMKLEVTYKKGCT--A 322
 QY 484 GOAISYVDGMQVYAIAG 502
 Db 323 GDPISWTSGLTAVTVIGG 341

RESULT 12

US-08-931-608A-5
 ; Sequence 5, Application US/08931608A
 ; Patent No. 6302685

GENERAL INFORMATION:

APPLICANT: Iobelt, Peter
 APPLICANT: Sleat, David E.
 TITLE OF INVENTION: NOVEL HUMAN LYSOSOMAL PROTEIN AND METHODS OF ITS USE
 NUMBER OF SEQUENCES: 12
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: David A. Jackson, Esq.
 STREET: 411 Hackensack Ave, Continental Plaza, 4th floor
 CITY: Hackensack
 STATE: New Jersey
 COUNTRY: USA
 ZIP: 07601

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/931,608A
 FILING DATE:

CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:

NAME: Jackson Esq., David A.
 REGISTRATION NUMBER: 26,742
 REFERENCE/DOCKET NUMBER: 601-1-077
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 201-487-5800
 TELEFAX: 201-343-1684
 INFORMATION FOR SEQ. ID NO: 5:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 635 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHETICAL: NO
 US-08-931-608A-5

Query Match 3.8%; Score 104.5; DB 4; Length 635;
 Best Local Similarity 21.1%; Pred. No. 0.4;
 Matches 121; Conservative 65; Mismatches 239; Indels 149; Gaps 28;

QY 1 MKPTSLMASAGALALIAAPAFQAVTPVTEDELLANPAGEMWISYQONDEYHSHSLTQIT 60
 Db 3 IEKTLFVALIALMSSLSAHEADAWSTHQAMSPASTQV-----LAASSTWTF 54
 QY 61 TENVGQLQVWANGMOPGKYQVTF---PLIHGVMLANPGVYQAIKAKTGDLIWEHRR 116
 Db 55 TGNAYTILMTGSPRIDGAATVLEADHPHVEVALKLRNP-DALQTF----- 100
 QY 117 QLPNIATLNS--FGE-----PFRGM--AIYANGVYASTQYSPFGCFVSGH 160
 Db 101 -LAGVTTPGSALFGKFLTFSQTERFPTQSOVDVAVHLOQAGFTNIEVAPNRLLSAD 159
 QY 161 DSATGEELMKNFYIPRAGEGDETWGNDYEFARMTGAMGO-----ITYDPTNLVHY- 212
 Db 160 GTGAATNGFRFSIKRFSANGREFFANDAPA-LVPASLGDVNMVNLGLQNSVAKHTLHHV 218
 QY 213 ----GSTAVGPASEFQRCPT-----GGLYGTNTPRAVRPDGELVWRHQT 255
 Db 219 YHPEDVTPGPVNGVQAANAHAHPDFAIYGGSSLPAAFTNPAV-----GITWGSIT- 273
 QY 256 PRDMDQECTFPMNVTNVDQPSMEGLQSLINPAATGERRLVGVPCKTGMQDPAE 315
 Db 274 -----QTVVDLNSFTSGAGLATVNSTITKVGSGTFRANDPSNGE-WSLDSQ 318
 QY 316 -----TGEFLMARDTNYNNIESTIDENGIYV---VNEDALIKELDEYDYCPTFLG 363
 Db 319 DIVIAGGVKQLIFYSANGSSSGITDAGITASYRRAVTDNIKILINSL----- 370
 QY 364 GRDMPASALN---PDSGIYFIPLNVCYDMMAVDOEFT---SMQYFN-TSNVTKLPQG 414
 Db 371 GEDETAAGQSGQAADALF-----QAVVAGQGFSTIASGDAGVYQSTPTSGSRP 422
 QY 415 --KDMIGRIDALIDISTGFTLMSVERAANYSP-VLSTGGVLFNGGTDYFRFALSQETGE 471
 Db 423 YVANSAGTV-KIDL---THYSVEPAS--SPYIVQVGGTLLTSGT-----TWSE 467
 QY 472 TLMOTRLATVASGOAISYVDGMQVYAIAGGGS 505
 Db 468 TVWNEGSLAIPASG-----DNNQRIMATGGGS 496

RESULT 13

US-08-928-361B-30
 ; Sequence 30, Application US/08928361B
 ; Patent No. 6071518

GENERAL INFORMATION:

APPLICANT: Petersen, Carolyn
 TITLE OF INVENTION: PEPTIDES, POLYPEPTIDES, GLYCOPROTEINS,
 TITLE OF INVENTION: THEIR FUNCTIONAL MUTANTS, VARIANTS ANALOGS AND FRAGMENTS
 FOR TREATMENT AND DETECTION/DIAGNOSIS OF CRYPTOSPORIDIUM
 NUMBER OF SEQUENCES: 30
 CORRESPONDENCE ADDRESS:

ADDRESSEE: PETERS, VERNY, JONES & BIKSA

STREET: 385 Sherman Avenue, Suite 6

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94306-1840

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:

Query Match	3.7%;	Score 101.5;	DB 3;	Length 1043;
Best Local Similarity	20.8%;	Pred. No. 1.7;		
Matches 138;	Conservative 75;	Mismatches 221;	Indels 231;	Gaps 38

Db 1188 SLNYPSEFNPQQIDEITGKPVDIVIGLFIDFSIGELIDIRANGLAOCYR-----

```

QY 270 VTN-----VDVQ--PSTEMEGLAS-----INP-----NATG-----E 295
Db 1248 ITTDEVTGLPDLTETGLPRDPVSGLPQLPNGTLVDPSSNKKPLIPGSHSGTNGTSGECSHE 1307
QY 296 RRVLTGVPCKTGT-MQOFD-----AETGEFLMARDTNY-----Q 328
Db 1308 KDPSTGKPLPMPNGLHPFDEDSGLINPBTGDKLQCSHSGTETMPVPKPKPGENGIMPE 1367
QY 329 NMIESIDENGIVTNE-----DAI-----LKELDVEYD-----V 357
Db 1368 QILEAL--NKLPTSNEVNISPRSSDAVDPDRPTNWNKISQTYQVDDKTTIPGSAASY 1425
QY 358 CPTFLG--GDMPSAAL--NPDGIYFIPLNNVCYDMAVDQ-----EFTSMDEVNTSNV 408
Db 1426 IHTALGPTQDPTTGLPSPSTGLPFIQFNVLPDQGEQIKSVPLVSLIYVEKVIY 1485
QY 409 TKLPGKDMIGRIDAIDISTGRIMVEGAANYSPLYSTG-----GCVLFNGCTDRYF 462
Db 1486 TEAYGL-----PDPPTGEPIPI-----SYLPFAKNGELIDPISGKYFSGS IAGFI 1533
QY 463 RA-----LSQETGFTLMQT-----RLATVAGSQAISYVDGMQYVALAG----- 501
Db 1534 SGKAGSQKSDSGSNPDPSTNMPYDKGKLLIDPSGIALDMSVSGV--FATVPGTAAP 1592
QY 502 --GGV 504
Db 1593 KKGCV 1597

RESULT 15
US-07-731-157A-4
Sequence 4, Application US/07731157A
GENERAL INFORMATION:
APPLICANT: Quax, Wilhelmus J.
APPLICANT: Misset, Onno
APPLICANT: Van der Laan, Jan M.
APPLICANT: Lenting, Herman B.M.
TITLE OF INVENTION: Mutated beta-lactam acylase genes
NUMBER OF SEQUENCES: 50
CORRESPONDENCE ADDRESS:
ADDRESS: COOLEY GODDARD CASTRO HUDDLESON & TATUM
STREET: FIVE PAUL ALTO SQUARE, 4TH FLOOR
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: PC-DOS/MS-DOS
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07731,157A
FILING DATE: 19910509
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: EP 90200962
FILING DATE: 18-APR-1990
ATTORNEY/AGENT INFORMATION:
NAME: RAE-VENTER PH.D., BARBARA
REGISTRATION NUMBER: 32,750
REFERENCE/DOCKET NUMBER: GBRO-027/00US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-494-7622
TELEFAX: 415-857-0663
TELEX: 380816 COOLEY PA
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 816 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear

```

```

MOLECULE TYPE: protein
US-07-731-157A-4
Query Match
Best Local Similarity 3.6%; Score 98.5; DB 1; Length 816;
Matches 102; Conservative 73; Mismatches 178; Indels 199; Gaps 30;

QY 40 EWISYQGNQENRHSPLTQITTEENVGQLQVWARGM-----QKQVQVPLHIDGMITLAN 95
Db 145 EYVDF-----DFQPEPLTDF-----DVMYVWGSANRFSPTNLEVALA----- 184
QY 96 PGYVIAIDAKTG-----DLTW-----EHRRLPNITLNSFGPT 131
Db 185 --WROSLKROHGPBRGRALFDELLIMINDTAPTTPAPAEHKKQQAQ-----AGT 232
QY 132 RGMATVANGVAVAGSTCOYSPGCFVSGHDSA--TGEELMRNYFIPRAGEGDETFWGDYE 190
Db 233 QDLAHVSSPVLATTELEROCKHMG--GRGPDAPKASNLIMSTR--PERVOEGSTVLINGPQ 288
QY 191 ARWMTGAMQOITYDPVTNLVHYGSTAVGPASETORGTGPG--TLXGNTFRFAVRPTGE 247
Db 289 FGV-----YNPAYT--YGIGLHAGAGFDVYGNTPFAIPVLFGTNS-----E 327
QY 248 IYWRHQTLPROMWDOECTFEMAVTNVDVQSTEMEGLQSPNNA-----T 293
Db 328 IAMGATAGPOD-----VVDIYQ-----EKLNPSSADQYMFNNMARTMEQR 367
QY 294 GRRVLTGVPCKTGTMM-----QPDATGE-----FLMARDTNYON 329
Db 368 KERIOVRCQADREMTITWTVHGPMQEDYDGAAYSKRSMWDGYEVSQSLAMLVAKARN 427
QY 330 MIESIDENG--IYVNEDEALIKEIDVEYDVCPTL-----GGRDMP-- 369
Db 428 WTEFLDQASKMAISINWYADKHGNIGY--VSPATLPQRPADQDIRPAKDGSMELGK 486
QY 370 -----AALNDSGIYFIPLNNVCYDMAVDQETSMDEVNTSNVTKLPKXDMIGRIDA 423
Db 487 SFDAIPKAYNPPQG-YLVNWN--KPAPDK--TNTDTYWTYGGRM--NELVSQYQO 536
QY 424 IDISTGRTMSEVERAANYSPLYSTGGCVLFNGCTDRYFRALSQETGLMOTRLATVAS 483
Db 537 KDLFSVQEIWFENQ-KASYSDV-----NMKTFRPHLEKLAQOLPADDSKAL 583
QY 484 GOAISYEVDGMQ 495
Db 584 TMLLAW--DGME 593

```

Search completed: May 24, 2002, 10:20:18
Job time: 331 sec

•
• Fri May 24 11:27:32 2002

walick-934-135.pep.ra1

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2002, 10:21:08 ; Search time 54.76 Seconds
(Without alignments)
898.425 Million cell updates/sec

Title: WALICK-934-135.PEP
Sequence: 1 MKPTSLMMSAGALALIAAP.....GMOYVAIAGGVSYGSLNS 512
Perfect score: 2742

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues
Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_71:*
1: PIR1:*
2: PIR2:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	495	18.1	742	2 A49340	alcohol dehydrogen
2	487	17.8	601	2 E95863	alcohol dehydrogen
3	485.5	17.7	708	2 S52317	quinohemoprotein e
4	483.5	17.6	738	2 S14270	alcohol dehydrogen
5	483	17.6	742	2 JS0326	alcohol dehydrogen
6	434.5	15.8	626	2 JQ0706	alcohol dehydrogen
7	417	15.2	623	2 B83399	methanol dehydroge
8	392.5	14.3	573	2 S68591	glucose dehydrogen
9	349.5	12.7	796	2 JY0107	glucose dehydrogen
10	349.5	12.7	796	2 H90644	glucose dehydrogen
11	349.5	12.7	796	2 H85495	glucose dehydrogen
12	340.5	12.4	796	2 AG0523	glucose dehydrogen
13	329	12.0	801	1 S00943	glucose dehydrogen
14	294.5	10.7	808	1 OPEKEX	glucose dehydrogen
15	279	10.2	803	2 F83360	glucose dehydrogen
16	276	10.1	809	2 B98314	glucose dehydrogen
17	276	10.1	809	2 A12968	glucose dehydrogen
18	268	9.8	778	2 G98221	glucose dehydrogen
19	268	9.8	778	2 A13064	glucose dehydrogen
20	254.5	9.3	809	2 A55547	quinone-shikimate
21	220	8.0	639	2 JC4881	polyvinyl-alcohol
22	209.5	7.6	524	2 A41378	hypothetical prote
23	157.5	5.7	524	2 A82580	polyvinylalcohol d
24	144.5	5.3	525	2 A99553	hypothetical prote
25	142.5	5.2	392	2 AG0821	probable lipoprote
26	142	5.2	839	2 D97013	probable secreted
27	140.5	5.1	668	2 C75264	probable serine/th
28	137.5	5.0	392	2 F91050	probable dehydroge
29	137.5	5.0	392	2 B85895	probable dehydroge

30	136.5	5.0	392	2 G65027	hypothetical prote
31	134.5	4.9	407	2 H69064	serine/threonine p
32	132.5	4.8	386	2 A82284	conserved hypothet
33	131	4.8	2535	2 AC0304	probable hemolysin
34	126	4.6	407	2 C82804	conserved hypothet
35	125.5	4.6	827	2 F64512	hypothetical prote
36	120.5	4.4	3573	2 S23070	erythropoietin synt
37	119.5	4.4	1588	2 A86036	probable adhesin B
38	119.5	4.4	1588	2 H91188	probable adhesin B
39	119.5	4.3	1526	2 AC2239	MD-40 repeat prote
40	118.5	4.3	799	2 T48889	serine/threonine p
41	117.5	4.3	1197	2 D82696	hypothetical prote
42	117	4.3	1354	2 AG0538	Rhs-family protein
43	115.5	4.2	954	2 A87431	regulatory protein
44	115	4.2	443	2 S23771	outer membrane por
45	114	4.2	796	2 T39962	coatomer complex b

ALIGNMENTS

RESULT 1
A49340 alcohol dehydrogenase (EC 1.1.-.-) precursor - Acetobacter pasteurianus (strain NC113
C:Species: Acetobacter pasteurianus
C:Date: 07-Apr-1994 #sequence_revision 07-Apr-1994 #text_change 20-Jun-2000
C:Accession: A49340
R:Takemura, H.; Kondo, K.; Horinouchi, S.; Beppu, T.
J. Bacteriol. 175, 6857-6866, 1993
A:Title: Induction by ethanol of alcohol dehydrogenase activity in Acetobacter pasteu
A:Reference number: A49340; MIMD:94042848
A:Accession: A49340
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-742 <TAK>
A:Cross-references: GB:D13893; MID:9517067; PIDN:BA440252.1; PID:9452586
C:Superfamily: glucose dehydrogenase (pyridoxinoline-quinone)
C:Keywords: alcohol metabolism; NAD; oxidoreductase

Query Match	18.1%	Score 495	DB 2	Length 742
Best Local Similarity	27.0%	Pred. No. 8.9e-28		
Matches 164	Conservative 82	Mismatches 221	Indels 140	Gaps 22
QY	11 AGATALLAAPAFQVTPYTDLL--ANPAGEWISYQONENYRHSPLQITTEYNGQLD 68			
DB	23 AAALFYAAVPPARAGOGGTGAIHADDHPENMLSYGRYSSEORYSPLDIINSNVDLK 82			
QY	69 IWMARGMPGRVQ-VTPLIHDGVMIANPGDVIOATDAKGLIWEHRQLP-NIA----- 122			
DB	83 LAMVYTLDTNRQENTPLVVDGIMVATTNMSKMEALDAATGKILMDYDPKVPGNADKGC 142			
QY	123 --TLN-----SFGK-----PTRGMAI 136			
DB	143 CDYVRGAGYNGKVFQWGTFGDGRVLAADAKTGKWEVNTIPADASIGKORSYVVDGAVR 202			
QY	137 VANGVIVAGSTOYSPFGC--FVSGHDSATGEBELRNFTIRAEEDD----- 182			
DB	203 VAKGLVLIGN--GGESEFGARFVSADFETGTLKWRFTVNNKNEPDHAAVDVNLMSKA 260			
QY	183 -ETWGDYDARAKMT-----GAMGQITVPTVNLVYSGTAVGPASETORGPGGTLV--- 233			
DB	261 YKTWGP--KGAVRQGGGTWDSLVYDPSDLIY---LAWG-----NGSPMYKVRSE 309			
QY	234 --GTN---TRFAVRPDTEIWRHQTLPRNDMDCEFFEMVTVNVQPTSEMDGLSI 287			
DB	310 GIOSNLFISIALVAKRETEGVVHFQATPMDQMDYTSVOQIWTFLDMPV----- 357			
QY	288 NPAAATGE--RRLVTPCCTGTMMQFDATGELNARDPTNYNMIESIDE--NGIYTVNED 345			
DB	358 -----NGEMRHVIMAP--KNGFTYVDAKTGFLAGKKNVYNNMANGDLPLGRIPYPD 411			
QY	346 AILKEIDVEYDVCPTFLGGRDMSAALNPDGIIYPLINNVCY-----DIMA 392			

Db 412 GLYTLTGKFWYGIPELGAHNFMGMAYSPTHTLVYLAHPIPCGYKNQVGFPHDPANN 471
 QY 393 VDOEFTSDVNTSVNWKLPKGMICRIDAIDISTGRTLSVERAANYSPVLTGGGV 452
 Db 472 VGLDMTKRGLDTPDE-ARTAVIKDLHEWLLAMPVKMEFYWKIDHGGPMNGVLTATGDL 530
 QY 453 LFNCGTDYRFRALSOETGETLMQTRLATVASGOAISYEVDGMQYVA-----IACG 502
 Db 531 LFOGLANGEFHAYDATNGSDLYKFKDAQSGIIASPMYTSVNGKOYVAEVEGGLPIPSMG 590
 QY 503 GVSYSG 509
 Db 591 GVGRTSG 597

RESULT 2

E95863
 Alcohol dehydrogenase (acceptor) (EC 1.1.99.8) large subunit [imported] - *Sinorhizobium*
 C:Species: *Sinorhizobium meliloti*
 C:Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 14-Sep-2001
 C:Accession: E95863
 R:Finan, T.M.; Weidner, S.; Wong, K.; Buhmester, J.; Chain, P.; Vorholter, F.J.; Hernan
 Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001
 A:Title: The complete sequence of the 1,683-kb pSymb megaplasmid from the N2-fixing endo
 A:Reference number: A95842; MUID:21396508; PMID:11481431
 A:Accession: E95863
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-601 <KUR>
 A:Cross-references: GB:AL591985; PIDN:CA648573.1; PID:915140045; GSPDB:GNO0167
 A:Experimental source: strain 1021, megaplasmid pSymb
 R:Galbert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abolia, P.; Ampe, F.; Barloy-Hubler,
 P.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;
 Science 293, 668-672, 2001
 A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,
 Hubbard, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.
 A:Title: The composite genome of the legume symbiont *Sinorhizobium meliloti*.
 A:Reference number: A96039; MUID:21368234; PMID:11474104
 A:Contents: annotation
 C:Genetics:
 A:Genome: Smb20173
 A:Genome: plasmid
 C:Keywords: oxidoreductase

Query Match 17.8%; Score 487; DB 2; Length 601;
 Best Local Similarity 26.2%; Pred. No. 2.5e-27;
 Matches 157; Conservative 90; Mismatches 215; Indels 138; Gaps 24;

QY 14 LALLAAPAFQVTPVDEL--LANPAGWISYSGONQENRHSPLTOITTEVGOQLVW 71
 Db 8 LAISIGGAGVAFANDELQKIDDP-NQMAITGDIYANIRSKLDQINKDNVGRKIQVAM 66
 QY 72 --ARGMQPKVQVTPPLIHGCVMTLANP-GDVIOAID-ARTGDLVEMHR-RQLNPATLINS 126
 Db 67 TFSFGVLRG-EGSPVLTGIDLVHTPPFNTVYALDLSKQGVYMKYKQDPNTPVWC 125
 QY 127 FGEPTRGMAI-----VANGYIVAGST----- 147
 Db 126 CDIVNGVAADAKIFLHQADITVVALDAKTKGVYMSVKNNGATGENTATVMPVKIKI 185
 QY 148 ---CQYSPGCG--FVSGHDSANGTEELMRNYFI-----PAGESEDPT 184
 Db 186 LVGISGEGVGRVHTAYAMADQKVLKRGYSKGPDSDTLIDPEKTHLGKPEVGDGSLTT 245
 QY 185 WGNDFYARMMTG--AMGQITDPVTLNLYHGSTAVGPASETFQRTGPGTLYGTNTPAV 241
 Db 246 WEED---QMKIGGTTGWTYSTDPEENLYYGTGNPSTWNPOR--PDGNBW-SWTIPAR 299
 QY 242 RPDGTETVHRHOTLPDNDMDQCTEMAVTNVDQPTSTEMGLSINMNAITGERVLTG 301

Db 300 DVDTGMAKWLQMTPHDEMDYDGVNEMILNEEQIDCK-----DRKLTH 343
 QY 302 VPCRTGMWQPDATGTEFL-----WARTNTQNMIE-----SIDENGIV 340
 Db 344 FD-RNGGTYNDRTVTEGLLVAEKYDPTVNNATEVVDPKSDKGRPOVQAQSTEQNG-- 400
 QY 341 TVNEDAILKEIDVEYDCPTFLGRDMPSSAALPDSGTYEFTPLNNVCYMAAVDEFTSM 400
 Db 401 ---EDTWT-----GCPAALGTCKDQOPRAYSRTKELTYFTNHCMDYEPFRVSYAG 451
 QY 401 DVYNTSNVTKLPKGMK---IGRIDAIDISTGRTLSVERAANYSPVLTGGVLFNGG 457
 Db 452 QPTVGAATLSWYP-KDSHGGMGNFIAMDNKEGKIKWSLPEPFSVSGALATAGDVVYCT 510
 QY 458 TDRYFRALSOETGETLMQTRLATVASGOAISYEVDGMQYVAAG-----GVSISGLNS 512
 Db 511 LEGYLAHVDAATCKELKXRTKTSVGNVMTTAREGKOYVAVLSGVGMAGITGLAAGLTN 570

RESULT 3

S52317
 quinolomoprotein ethanol dehydrogenase (EC 1.1.99.-) type 1 precursor - *Comamonas tes*
 C:Species: *Comamonas testosteroni*
 C:Date: 08-May-1995 #sequence_revision 21-Jul-1995 #text_change 02-Jun-2000
 C:Accession: S62366; S62373; S65908; S52317
 R:Stoorvogel, J.; Knaayveld, D.E.; van Suijs, C.A.; Jongejan, J.A.; de Vries, S.; Dui
 Eur. J. Biochem. 235, 690-698, 1996
 A:Title: Characterization of the gene encoding quinolomoprotein ethanol dehydrogenas
 A:Reference number: S62366; MUID:96184549
 A:Accession: S62366
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-708 <STO1>
 A:Cross-references: EMBL:X81880; NID:9663195; PIDN:CAA57464.1; PID:9663196
 A:Accession: S62373
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 32-54 <STO2>
 R:de Jong, G.A.H.; Geerlof, A.; Stoorvogel, J.; Jongejan, J.A.; de Vries, S.; Dui
 Eur. J. Biochem. 230, 899-905, 1995
 A:Title: Quinolomoprotein ethanol dehydrogenase from *Comamonas testosteroni*. Purific
 A:Reference number: S65908; MUID:95324580
 A:Accession: S65908
 A:Molecule type: protein
 A:Residues: 32-50; 'X', 52-54; 477-483, 'X', 485-490 <DEU>
 A:Experimental source: ATCC 15667
 C:Genetics:
 A:Gene: qhedh
 C:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)
 C:Keywords: oxidoreductase; quinolomoprotein
 F:1-31/Domain: signal sequence #status predicted <SIG>
 F:32-708/Product: quinolomoprotein ethanol dehydrogenase type 1 #status experimental

Query Match 17.7%; Score 485.5; DB 2; Length 708;
 Best Local Similarity 25.7%; Pred. No. 4e-27;
 Matches 152; Conservative 86; Mismatches 241; Indels 113; Gaps 20;

QY 3 PSLIMASAGALALLAAPAVQVTPY-----DELANP-AGENISYSGONEN 50
 Db 12 PGRMWTLLAACLG--SAAPFAQTGPAAQAAAVQVDDDFRANARPPDPTIGVYAE 69
 QY 51 YRHSPLTQITTEVGOQLVMAKQMPKQ-VQVTPPLIHGCVWYLANPQDVIOAIDARTGD 109
 Db 70 TRYSRLDQIANANVKDLGLAWSYNLESTRGVATFPVVDGLIMYSASVYHAIDRTGN 129
 QY 110 LIMEHRROP-----NIAT-----LNSFG 128
 Db 130 RIWITDPQIDRSTGKCGCDVVRGVALMKGVYVGAWDGLIALDAATGKEVHWQTF- 188
 QY 129 EPTRGMAIYVANG-----VIAGSTCOYSPFGCFVSGHDSANGTEELMRNYFI----- 175
 Db 189 EGQKSLFITTGAPRVFKKGVILIKRGALEYGRG--YITAYDAETGERKRWFSVPGDSKP 247

QY 288 NPNATGE-RHYLGVPGCKTGMQFPAETGEFLMARDNYOMNISID-----BN 337
 Db 359 -----GEMRHYIYHAP-KNGFFVLDLAKTEGELSGNIYYQWANGCLDPLTGRPMYNDP 411
 QY 338 GIYVNEALIKELDEYDVCPTFLGGRDWPSSAALNDGSIYIPLNNVCY-----388
 Db 412 GLYTLNG-----KFWYGI-PGFLGAHNFMAAAYSPKTHLYIIPAHQIPFGYKNGVGGE 463
 QY 389 -----DMAVDQFTSMQVNTSVNTKLPKGDMDIGRIDALDISTGRTLSVERAANYSF 444
 Db 464 KPHADSNVNGVDMTKKNLPTPE-ARKAYIKDLHGWLMLNDPVKMETWKIDHKGPMNGC 522
 QY 445 VLSGGCVLNGGTDRYFRALSOETGETLMQTRLATVASGOAISYEDMOQYVA-----498
 Db 523 ILATGDLLEFQGLANGEFHAYDATNGSDLYKFDAGSIIAPPMYTSVNGKQYVAIVEVNG 582
 QY 499 -----IAGGVSYSYG 509
 Db 583 GIYPSMGVGRITSG 597

RESULT 6

JQ0706
 alcohol dehydrogenase (acceptor) (EC 1.1.99.8) alpha chain precursor - Methylobacterium
 N:Alternate names: methanol dehydrogenase 62K large chain
 C:Species: Methylobacterium extorquens
 C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 08-Oct-1999
 R:Accession: JQ0706; S07908
 R:Anderson, D.J.; Morris, C.J.; Nunn, D.N.; Anthony, C.; Lidstrom, M.E.
 Gene 90, 173-176, 1990
 A:Title: Nucleotide sequence of the Methylobacterium extorquens AM1 moxP and moxJ genes
 A:Reference number: JQ0706; MUID:90337342
 A:Accession: JQ0706
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-626 <AND>
 A:Cross-references: GB:M31108; NID:g150017; PIDN:AAA2380.1; PID:g150018
 A:Experimental source: strain AM1
 R:Nunn, D.N.; Day, D.; Anthony, C.
 Biochem. J. 260, 857-862, 1989
 A:Title: The second subunit of methanol dehydrogenase of Methylobacterium extorquens AM1
 A:Reference number: S04644; MUID:89350892
 A:Accession: S07908
 A:Molecule type: protein
 A:Residues: 28-50, 'XX', 53 <NUN>
 A>Note: The source is designated as Methylobacterium extorquens AM1
 C:Comment: This enzyme oxidizes methanol to formaldehyde.
 C:Genetics:
 A:Gene: moxP
 C:Keywords: alcohol metabolism; oxidoreductase
 F:1-27/Domain: signal sequence #status predicted <SIG>
 F:28-626/Product: alcohol dehydrogenase (acceptor) #status experimental <MAT>

Query Match 15.88; Score 434.5; DB 2; Length 626;
 Best Local Similarity 26.34; Pred. No. 1.6e-23;
 Matches 156; Conservative 91; Mismatches 219; Indels 127; Gaps 26;
 QY 10 SAGALALLA-AP-AFAQVTPVTELLANPAGE-WISYONQENYHSHPLTOITENVGQ 66
 Db 7 SVSLALMALALPALSSGAYANLXKSLSDNWMVPGKNDSDNMFSLDKOIKNGNVKQ 66
 QY 67 LQLVN-ARGMOPGVQVPLIHGVN-----LANPGDY-----100
 Db 67 LRPAWTFSTGLNGH-EGAPLVVDGKMYIHSPNNTPALGLDDETTIMQKPKQNPAA 125
 QY 101 QAT-----DAKTGDLIWEHRRLP-NIATLNSGFEPTRGMAIVANGVIYA 144
 Db 126 RAACACDLVNRGLAYWPGDKTPLLIL-KTQLDGNVALNA-ETGEIVMKVEHSDIKV 181
 QY 145 GSTCYSPF-----GC-----FVSGDSATGELMKNYFI-----174
 Db 182 GSTLTITAPYVVKKVIIGSSGAEGLGVRYLTAVDYKTBQVRAVATGPDKDLIASDFN 241

QY 175 ---PAGEE--GDETMQNDYEAHMTCG---WQOITYDPVTLNVHGSTAVGAPASETORG 226
 Db 242 IKNPYQKGLGCTWEGD---AMKIGGTNMGWYAYDETNLIFGTGNPAPMNETMK- 297
 QY 227 TPQGTILGTNTRFVAPDPTGEIYWRHQLTPRDMWDECTEFEMVTVNVDQSTEMEGLQS 286
 Db 298 -PGDNKM-TWYIIGRDAIDTGEAFKGYOKTPHDEMD-----YAGVNNMLSECKD--- 344
 QY 287 INPAAATGERRYLVGVPGCKTGMQFPAETGEFLMAR---DTNYOMNISIDENGIYVN 343
 Db 345 -----KDGARKLLTHPDNRNGIYVTLDRDGLVSNKRLDPT--VNVKSDPLKTGPYR 397
 QY 344 EDALIKELD-VEYDVCPTFLGGRDWPSSAALNDGSIYIPLNNVCYDMAVADQFTSMQV 402
 Db 398 DPEYGTMDHLADICPSAAGYHNGHSDYDPKRELFEFGNIHICMDPEPMLPRAQF 457
 QY 403 YNTSNYTKLPPCK-----DMIGRIDALDISTGRTLSVERAANYSFVLSGGCVLFG 456
 Db 458 FVGATLMMTPGPKGRONTGEGQIKAYNAITGDYKWKEMERFAVWGSTMATAGDLVFG 517
 QY 457 GTDRYFRALSOETGETLMQTRLATVASGOAISYEDMOQYVAIAGGVSYSYG 509
 Db 518 TLDGYLKARDSDPTGDLMLKFKIPSGALIGYPMYTHKGTQYVAI-----YIGVG 565

RESULT 7

B83399
 quonoprotein alcohol dehydrogenase PA1982 [Imported] - Pseudomonas aeruginosa (strain
 C:Species: Pseudomonas aeruginosa
 C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
 C:Accession: B83399
 R:Stoyer, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.;
 adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; L
 Nature 406, 959-964, 2000
 A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pa
 A:Reference number: A82950; MUID:20437337
 A:Accession: B83399
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-623 <STO>
 A:Cross-references: GB:AE004624; GB:AE004091; NID:g9947973; PIDN:AA05370.1; GSPDB:GN
 A:Experimental source: strain PA01
 C:Genetics:
 A:Gene: exaA; PA1982

Query Match 15.24; Score 417; DB 2; Length 623;
 Best Local Similarity 23.18; Pred. No. 3e-22;
 Matches 140; Conservative 95; Mismatches 247; Indels 124; Gaps 20;
 QY 3 PTLMSAGALAL--LAAPAFQVTPVTELLAN--PPAGEWISYONQENYHSHPLT 57
 Db 9 PAGLIRPSLHCLAAVAVALGSAGAAALADVYMEDIANDDTKTGVDLOYGCTGAQWSPK 68
 QY 58 QITTEYNGOLQVWARGM---OPGKQVPLIHDGVMIANGVIOIADKTTGDLTME 113
 Db 69 QVADNVETKLPMSYSFGEDEKRGQ-ESQATYSDEVITYTYSKRLPALDKTKRLMT 127
 QY 114 HRRLP-----NIATLNS-----126
 Db 128 YNHRLPDIPCCDVVNRGAATYDKVFEGTLDAVALNKWTGKVVYMKKRPADHAGAYT 187
 QY 127 -FGEPTGMAIVANG-----VIYAGSTQCYSPGCE--VSGDSATGELMKNYFI-- 176
 Db 188 MTGAPF-----IYADGKTGKVLTHGSS--GDEFVVGRLFAFADPDTGEBIWRPVEGSHM 241
 QY 177 -----AGEGDETMQNDYEA-----RWMTGA---WQOITYDPVTLNVHGSTAV 217
 Db 242 GRUNGKDSYTGDKVAKPAMPDDRNSTPKVESHSKGGAPWOSAFDEMTITIIYVAGKNP 301
 QY 218 GPASEYQKGTGPG-----TLVGTNTRFAVRPDPTGEIYWRHQLTPRDMWDECTEFEMVTV 271

```

Db 302 GPNWTAARAKAGNDHSDSLX-TSGQYGVDPSSSEVWFQHPNPDAWPFSGNNELVLF 360
OY 272 MNDVDPSTMEGLQSTINP-----AATGERVLYGVPCCKITGM--QEDAFETGFLWARD 324
Db 361 DYAKAGKIVKATATAHADRNGFFVYVDRSNGKIQNAFPVDNITWASHIDLKGTG-----RP 415
OY 325 TYNQNMIESIDENGLVTWEDALIKELDEYEVCPTEFLGHDWPSAALNPDGITYEPLN 384
Db 416 VEREGORPPLPEPG-----QKHKAYVSPPLFGCKNWNMAYSQDTGLFYPVAN 465
OY 385 NVCYDMAVDOEFTSMVYNTSNVTKLPCKDMIGRIDALIDISTRTTWSERAANSP 444
Db 466 HMKEDVTEEVSYTKGSAVLGCMGFRIKRMVDHVGSLRAMPVSGKYWEKHEHLPWAG 525
OY 445 VLTGGGVLENGTDTYFRALDSQETGELMOTRLAVASGAISYEYGMQVYALAGGV 504
Db 526 VLTAGLVLTGTGIDGFFRAFDKSKELMKPOTGSGIVSPITWEOGEOYLGVT---V 582
OY 505 SYGSGL 510
Db 583 GYGAV 588

```

```

RESULT 8
S68591
methanol dehydrogenase (EC 1.1.1.244) heavy chain - Methylotrophus methylotrophus (strain
C:Species: Methylotrophus methylotrophus
A:Variety: strain W31
C:Date: 06-Dec-1996 #sequence_revision 13-Mar-1997 #text_change 08-Oct-1999
C:Accession: S68591
R: Xia, Z.; Dai, W.; Zhang, Y.; White, S.A.; Boyd, G.D.; Mathews, F.S.
J. Mol. Biol. 259, 480-501, 1996
A:Title: Determination of the gene sequence and the three-dimensional structure at 2.4 Å
A:Reference number: S68591; MUID:96256524
A:Accession: S68591
A:Status: preliminary; nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-573 <X1A>
A:Cross-references: EMBL:U01040; NID:g1127819; PIDN:AA83765.1; PID:g1127820
A:Note: the authors did not translate the codons for residues 1-2
C:Keywords: oxidoreductase

```

```

Query Match 14.3% Score 392.5; DB 2; Length 573;
Best Local Similarity 24.6% Pred. No 1.6e-20;
Matches 136; Conservative 84; Mismatches 218; Indels 115; Gaps 22;
OY 38 AGEMISYSGNDENYRHSPLTQITENNQQLVW--ARGMQPKVQVTPLIHDCVMTL-- 93
Db 13 AGAWPLATGGYYSQNSPLAQINKSVNKAAMSFSTGVNGH-EGAPLIVIGDMNVHS 71
OY 94 ANEGVIOAIDAKTDLLWEHR-RQLPNIALNLSGFEETRGMA-----IV---ANGVYA 144
Db 72 AEPNNYVALNLNDPKGIWQHKPKDASTKAVMCCDVVDKGLAIGAGQIVKAKANGHLLA 131
OY 145 -----GSTQYSPF-----GC-----FVSGHDSATGEBL 168
Db 132 LDAKTKIMWEYEVCDPKKGSLLTQAPFAKQTVLMGSGAELGVRAVNAFDTKTEBLK 191
OY 169 WR-----NYFLPRAGE--EGDETWGNDYEARMWGA---WGQITVDPVT 207
Db 192 WRAFATGSDSVRLAKDFNSANPHYGQFLGRTWEGD---AMKIGGSTNGWATYADPKL 248
OY 208 NLYHVSSTAVGPASETQRTGG-----TLVGTNTRFAVRPDTGEIVRHQTLPRDNMQ 262
Db 249 NLYYSGNAPANWETMR--PGDNKMTIINWRDL-----DTGAKKGYQKTPHDEWD 300
OY 263 ECTFEMAVTVNDVQSTMEGLQSTINPNAATGERVLYGVPCCKITGMQFDAETGEFLWA 322
Db 301 AGVQWAVLTD-----QPVN-----GKMTPLSHIDRNGLITVLTNRENGNLIVA 343
OY 323 RDTN-YQNMIESIDENGLVTWEDALIKELDEY-VDVCPTEFLGHDWPSAALNPDGITYE 380

```

```

Db 344 EKVDAVAVVEKVKDLTKGTGPVRDPEFATRMHKGNTICPSAMGFHNOGVDSYDPESTRTY 403
OY 381 IFLNNVCYDMAVDOEFTSMVYNTSNVTKLP-----PEKDMIGRIDALIDISTGRTLMSVE 436
Db 404 AGLINICMDMEPFMLPYRAGQFVGATLAMYPGPNGLPKKMGQIRAFDLTTGAKAKTKW 463
OY 437 RAANASPYLSTGGGVLENGTDTYFRALDSQETGELMOTRLAVASGAISYEYDGMQY 496
Db 464 EKFAWAGTLTYKGLVWATIDYGLKALDNKDKGELNFMKMPGSGIGSPMTYSPFKQY 523
OY 497 VALAGGVSYSGG 509
Db 524 I-----GSMYGVG 531

```

```

RESULT 9
JV0107
glucose dehydrogenase (pyrroloquinoline-quinone) (EC 1.1.99.17) - Escherichia coli
C:Species: Escherichia coli
C:Date: 30-Jun-1992 #sequence_revision 12-Dec-1997 #text_change 16-Jun-2000
C:Accession: D64735; JV0107; A45997; S45201; I41228
R: Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.;
A.: Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID:97426617
A:Accession: D64735
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-796 <BLAT>
A:Cross-references: GB:AE000122; GB:U00096; NID:g1786315; PIDN:AAC73235.1; PID:g17863
A:Experimental source: strain K-12, substrain MG1655
R: Cleton-Janssen, A.M.; Goosen, N.; Fayet, O.; van de Putte, P.
J. Bacteriol. 172, 6308-6315, 1990
A:Title: Cloning, mapping, and sequencing of the gene encoding Escherichia coli quino
A:Reference number: JV0107; MUID:91035240
A:Accession: JV0107
A:Molecule type: DNA
A:Residues: 1-58, 'L', 60-148, 'H', 150, 'KRRCHT', 157-192, 'K', 194-796 <CLE>
A:Cross-references: GB:X51323; NID:941553; PIDN:CAA35706.1; PID:g41554
A:Experimental source: strain K12
R: Yamada, M.; Sumi, K.; Matsushita, K.; Adachi, O.; Yamada, Y.
J. Biol. Chem. 268, 12812-12817, 1993
A:Title: Topological analysis of quino protein glucose dehydrogenase in Escherichia co
A:Reference number: A45997; MUID:93286127
A:Accession: A45997
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-20 <YAM>
R: Fujita, N.
submitted to the EMBL Data Library, January 1994
A:Reference number: S45181
A:Accession: S45201
A:Molecule type: DNA
A:Residues: 1-796 <FUJ>
A:Cross-references: EMBL:D26562; NID:9473770; PIDN:BA00580.1; PID:g473791
A:Experimental source: strain K-12 substrain W3110
R: Tamada, M.; Asaka, S.; Sailer, M.H.; Yamada, Y.
J. Bacteriol. 175, 568-571, 1993
A:Title: Characterization of the gsd gene from Escherichia coli K-12 W3110 and regula
A:Reference number: I41228; MUID:95123180
A:Accession: I41228
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-148, 'H', 150, 'KRRCHT', 157-192, 'K', 194-665, 'H', 667-796 <RES>
A:Cross-references: GB:D12651; NID:921655; PIDN:BA002174.1; PID:g216556
C:Genetics:
A:Gene: gsd
A:Map position: 3 min
C:Function:
A:Description: catalyzes the oxidation of D-glucose to D-gluconic acid by ubiquinone
A:Pathway: respiratory chain

```

C:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)
 C:Keywords: oxidoreductase; pyrroloquinoline quinone; respiratory chain; transmembrane
 F:11-37/Domain: transmembrane #status predicted <TM1>
 F:41-59/Domain: transmembrane #status predicted <TM2>
 F:63-81/Domain: transmembrane #status predicted <TM3>
 F:96-110/Domain: transmembrane #status predicted <TM4>
 F:120-140/Domain: transmembrane #status predicted <TM5>
 F:93-95/Binding site: ubiquinone (Arg, Asp) #status predicted
 F:466/Active site: Asp #status predicted

Query Match 12.7%; Score 349.5; DB 1; Length 796;
 Best Local Similarity 22.8%; Pred. No. 3.3e-17;

Matches 161; Conservative 73; Mismatches 209; Indels 263; Gaps 30;

```

QY 6 LMAS-----AGALALLAFAQVPTVDLLANPPAGEWISYGOENYRHSPLTQ 58
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 135 LTMAGFNDOEINGTSLADATPAEA-ISPVDQ-----DMPAYGRNOCGRFSPKQ 185

QY 59 ITTENVGOLQVWA-----RGMPGKV--QVPTLHDGVMYLANPGDVIQAIIDAKTD 109
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 186 INADVHNKEMAVFRTGDVQKQPNDEGITTNEVPIKVGDTLYICTAHORLFDALDASGK 245

QY 110 LWEHRRLPNTATLNSFG-----EPTRGMAIVANGVIYAGS--TCQYS--PFG 154
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 246 EKMHYD---PELKTNEFQHTVCRGVSYHEAKAETASPEVADCPRIILPVDGRLLAI 302

QY 129 -----EPTRGMAIVANGVIYAGS--TCQYS--PFG 154
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 303 NAENKLCETFAKGVNLQSNMPTDKPGLYEPTSPRIITDKTIVMAGSYTDNFTSTETS 362

QY 155 CEVSGHDSATGEBELRNFEIPRA-----GEEDDETGNDEYARMTGAMQITVDPT 207
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 363 GVIRFEDVNTGELMA--FDGADPNAPISDEHTFTENS-----PNSMAPAADAAL 413

QY 208 NLVHGSTAVGPASSETQGTGTLGNTNFAVRPTDGTGIVRROTLPDNDMDQECTFE 267
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 414 DLVYLPKGVTPPDLMGKRNTPDEQERIASI--LALNATTGKLAWSYOTVHNDLMDMD---- 468

QY 268 MMVTNVDPSTENEGLOSINPNAATGERVLTGVPCKTGMQFADTEGEFL----- 320
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 469 ---LPAQPT---LADITVN--GQKVPVIYAPAKTGINIFVLDNRNGELVVAPEK 515

QY 321 ---WARDTNYQNMIESIDENGI-----VTVNEDALIKEL--DVEYD----- 356
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 516 VPQGAAGDYVTPQPSSELSFRPTKDLGADMMGATMFDOLVCRYMFHOMRYEGITFP 575

QY 357 -----VCPFTLGRDMPSPAALNP----- 374
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 576 SEQGLTVEFPGLMFEWGGISVDPNREVAIANPMALPFVSKLIPRGPNMEOPRKAKGT 635

QY 375 ---DSGI-----YFIPLNVCYDMAVADQETSMQVNTSVNTKLPCKDMIGRIDAID 425
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 636 GTESGIQPOYGVPTVTLN-----PFLS-----PFGLPCKOPAMGYISALD 676

QY 426 ISTGRTLSVERAANYS-----PV-----LSTGGVLEFNGGT--DRTYRAL 465
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 677 LKTNEVVMKKRIGTPDQSMFPMPVPVPPNMGMPMLGCPISITAGNVLEITAAADNYLRA 736

QY 466 SOETGETLQTRLATVYASGA--ISEVDGMQYAI--AGGVSYSGS 508
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 737 NMSGKILWGRLP--AGGOATPMTEVNGKQYVVISAGHGSFST 780

```

RESULT 10
 H90644
 glucose dehydrogenase [imported] - Escherichia coli (strain O157:H7, substrain RMD 0509)

C:Species: Escherichia coli
 C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Aug-2001
 C:Accession: H90644
 R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
 gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
 DNA Res. 8, 11-22, 2001

A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and 9
 A:Reference number: A99629; MID:21156231; PMID:11258796
 A:Accession: H90644
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-796 <HAY>
 A:Cross-references: GB:BA000007; PIDN:BA93551.1; PID:q13359584; GSPDB:GN00154
 A:Experimental source: strain O157:H7, substrain RMD 0509952
 C:Genetics:
 A:Gene: Ecs0128
 C:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)

Query Match 12.7%; Score 349.5; DB 2; Length 796;
 Best Local Similarity 22.8%; Pred. No. 3.3e-17;

Matches 161; Conservative 73; Mismatches 209; Indels 263; Gaps 30;

```

QY 6 LMAS-----AGALALLAFAQVPTVDLLANPPAGEWISYGOENYRHSPLTQ 58
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 135 LTMAGFNDOEINGTSLADATPAEA-ISPVDQ-----DMPAYGRNOCGRFSPKQ 185

QY 59 ITTENVGOLQVWA-----RGMPGKV--QVPTLHDGVMYLANPGDVIQAIIDAKTD 109
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 186 INADVHNKEMAVFRTGDVQKQPNDEGITTNEVPIKVGDTLYICTAHORLFDALDASGK 245

QY 110 LWEHRRLPNTATLNSFG-----EPTRGMAIVANGVIYAGS--TCQYS--PFG 154
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 246 EKMHYD---PELKTNEFQHTVCRGVSYHEAKAETASPEVADCPRIILPVDGRLLAI 302

QY 129 -----EPTRGMAIVANGVIYAGS--TCQYS--PFG 154
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 303 NAENKLCETFAKGVNLQSNMPTDKPGLYEPTSPRIITDKTIVMAGSYTDNFTSTETS 362

QY 155 CEVSGHDSATGEBELRNFEIPRA-----GEEDDETGNDEYARMTGAMQITVDPT 207
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 363 GVIRFEDVNTGELMA--FDGADPNAPISDEHTFTENS-----PNSMAPAADAAL 413

QY 208 NLVHGSTAVGPASSETQGTGTLGNTNFAVRPTDGTGIVRROTLPDNDMDQECTFE 267
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 414 DLVYLPKGVTPPDLMGKRNTPDEQERIASI--LALNATTGKLAWSYOTVHNDLMDMD---- 468

QY 268 MMVTNVDPSTENEGLOSINPNAATGERVLTGVPCKTGMQFADTEGEFL----- 320
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 469 ---LPAQPT---LADITVN--GQKVPVIYAPAKTGINIFVLDNRNGELVVAPEK 515

QY 321 ---WARDTNYQNMIESIDENGI-----VTVNEDALIKEL--DVEYD----- 356
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 516 VPQGAAGDYVTPQPSSELSFRPTKDLGADMMGATMFDOLVCRYMFHOMRYEGITFP 575

QY 357 -----VCPFTLGRDMPSPAALNP----- 374
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 576 SEQGLTVEFPGLMFEWGGISVDPNREVAIANPMALPFVSKLIPRGPNMEOPRKAKGT 635

QY 375 ---DSGI-----YFIPLNVCYDMAVADQETSMQVNTSVNTKLPCKDMIGRIDAID 425
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 636 GTESGIQPOYGVPTVTLN-----PFLS-----PFGLPCKOPAMGYISALD 676

QY 426 ISTGRTLSVERAANYS-----PV-----LSTGGVLEFNGGT--DRTYRAL 465
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 677 LKTNEVVMKKRIGTPDQSMFPMPVPVPPNMGMPMLGCPISITAGNVLEITAAADNYLRA 736

QY 466 SOETGETLQTRLATVYASGA--ISEVDGMQYAI--AGGVSYSGS 508
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 737 NMSGKILWGRLP--AGGOATPMTEVNGKQYVVISAGHGSFST 780

```

RESULT 11
 H85495

glucose dehydrogenase [imported] - Escherichia coli (strain O157:H7, substrain EDL933)
 C:Species: Escherichia coli
 C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
 C:Accession: H85495
 R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; May

iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamoculis, K.; Apodaca, Nature 409, 529-533, 2001
 A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
 A:Reference number: AB5480; MUID:21074935; PMID:11206551
 A:Accession: H85495
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-796 <STO>
 A:Cross-references: GB:AE005174; NID:912512839; PIDN:AA654428.1; GSPDB:GN00145; UMGCP:201
 A:Experimental source: strain O157:H7, substrain EDL933
 C:Genetics:
 A:Gene: gcd
 C:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)

Query Match 12.7%; Score 349.5; DB 2; Length 796;
 Best Local Similarity 22.8%; Pred. No. 3.3e-17;
 Matches 161; Conservative 73; Mismatches 209; Indels 263; Gaps 30;

6 LMAS-----AGALALIAAPFAQVTPVDELAPPAEEMISYQONENRHSPLTQ 58
 135 LTMAGFNDPQETNGTLISADATPAEA-ISPVAQ-----DWPAYGRNGEQGRSPKQ 185
 59 ITTENYGOLOLYWA-----RGMPGRV--OVTPLHDGYMTLANPQDVIQALDAKTGD 109
 186 ITADNVHNLKEMAVFRTGDKQPNPDEITNEVTIKVGDLYLCTAHQRLFALDAASGK 245
 110 LIMEHRROLPIATLNSFG----- 128
 246 EKHNYD--PELKTNESFOHYTCRGVSYHEAKATASPEVMADEPRRIILPVNDGRILAI 302
 129 -----BPTRGMAIVANGVIVAGS-TCQYS--PFG 154
 303 NAENGKLCETFAKGVNLQSNMPTKPGLYEPTSPRIITDKTIYMASSVDNESTRETS 362
 155 CFSGHDSATGEELMKNYFIIPRA-----GEGETWNGDYEARMTGAMQOITYDYVT 207
 363 GVRIGFVNTGELMA--FDPGAKDPAIPSEHTFTFNS-----PNSMAPAAYDKL 413
 208 NLVHYGSTANGPASETORGTGTLGCTNRFAVRNEDTEIWMRQTLPRDMDOECTFE 267
 414 DLVYLPMGVTPPIWGNRTPBOERYAST--LALNATTKLAMSQTVVHDLMDMD----- 468
 268 MMVTNVVOPSTEMEGLOSINPAATGERRVLTGVPCKTGTMOQDAETGEFL----- 320
 469 -----LPAQPT-----LADITVN--GQKVPYIAPAKGNIFVDRBRGELVVAPEKP 515
 321 ---WARDNTQNMIESIDENGI-----VTNEDALIKEL--DVEYD----- 356
 516 VPGAAGADVYTPIQFSELSFRTKDLGADMMGATMEDQLCRVFHMOREGIEFTPP 575
 357 -----VCPFTLGRDMPSAALNP----- 374
 576 SEQGTIVFPNGLMGEWGGISVDPNREVALIANPMALPFVSKLIPRGPNMDEPKAKGT 635
 375 --DSGI-----YFIPUNVVCYDMAVVDQETSMVYNTSVTKLPKGDIMIGRIDALD 425
 636 GTESSIQOQYGVPGVTLN-----PFLS-----PFGJPCCKOPANGYISAD 676
 426 ISTGRTIMSVRAAANTS-----PV-----ISTGGVLFNGGT--DRYFRAL 465
 677 LKTNVVWKKRIGTPQDSMPMPVPFPMFGMPLGCGISTAGNVLFIAATADNTLRAY 736
 466 SEQETLMQTRLATVVASGA--ISYEVDGMOYVAI--AGGVSYS 508
 737 NMSNGEKLMOGRLP--AGGQATPMFTEVNGKQYVVISAGGHSFGT 780

RESULT 12
 AG0523
 glucose dehydrogenase [imported] - Salmonella enterica subsp. enterica serovar Typhi (st
 C:Species: Salmonella enterica subsp. enterica serovar Typhi
 A:Note: this species has also been called Salmonella typhi

C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #ext_change 27-Nov-2001
 C:Accession: AG0523
 R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Main, J.; Church
 th, T.; Connerton, P.; Croplin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farr
 , S.; Moutle, S.; O'Gea, P.
 Nature 413, 848-852, 2001
 A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmons, M.; Skellon, J.; Stevens,
 A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica se
 A:Reference number: AB0502; PMID:11677608
 A:Accession: AG0523
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-796 <PAR>
 A:Cross-references: GB:AL513382; PIDN:CAD01327.1; PID:gl6501455; GSPDB:GN00176
 C:Genetics:
 A:Gene: STY0191
 C:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)

Query Match 12.4%; Score 340.5; DB 2; Length 796;
 Best Local Similarity 23.3%; Pred. No. 1.5e-16;
 Matches 160; Conservative 66; Mismatches 235; Indels 227; Gaps 29;

6 LMAS-----AGALALIAAPFAQVTPVDELAPPAEEMISYQONENRHSPLTQ 58
 135 LTMAGFNDPQETNGTLRADATPA--ATSSIADE-----DWPAYGRNGEQGRSPKQ 185
 59 ITTENYGOLOLYWA-----RGMPGRV--OVTPLHDGYMTLANPQDVIQALDAKTGD 109
 186 ITADNVHNLKEMAVFRTGDKQPNPDEITNEVTIKVGDLYLCTAHQRLFALDAASGK 245
 110 LIMEHRROL----- 118
 246 EKHNFDPQKLTSSFOHYTCRGVSYHEAKADTASPEVIADCSRILIPVNDGRLEAVNAE 305
 119 -----PIATLN-----SFEPTRGMAIVANGVIVAGS-TCQYS--PFGFV 157
 306 TGKLCETFAKGVNLQTNMPTTGLYEPTSPRIITDKTIYMASSVDNESTRETS 365
 158 SGHDSATGEELMKNYFIIPRAGEGDETNQNDYEARMT--GAMQOITYDYVNLVHYGST 215
 366 RGFVNSGKLMA--FDP--GAKDPTIIPADEHAFTFNSPNSMAPAAYDAKIDLVLPMG 421
 216 AVGPASETORGTGTLGCTNRFAVRNEDTEIWMRQTLPRDMDOECTFEEMVTVNDV 275
 422 VTPEDITNGKRTPEOERYAST--LALNATTKLAMSQTVVHDLMDMDLPAQPTLADIV 480
 276 QPSTEMEGLOSINPAATGERRVLTGVPCKTGTMOQDAETGEF----- 319
 481 DGLTV-----PVIAAPAKTGNIFVDRBRGELVVAPEKVPQGAAG 523
 320 -----LNAEDNTQNM-----MIESIDENGI--VTNEDALIK 349
 524 DYAKTOPFSDLTFRPKKDLGADMMGA--TMFQOLCRVWFHQLRREGIFTPPSEGTL- 581
 350 ELDEYDVCPFTLGRDMPSAALNPDGIFY-----IPUNVVCY-----DMA 392
 582 -----VFPNGLMGEWGGISVDPDROVALIANPMALPFVSKLIPRGPNMDEPKAKGT 634
 393 VDQETSMVYNTS--NVY-----KLPKGMIGRIDALDISTGRTIMSVRAAANTS 443
 635 TGTEAGIOPQYGVPGVTLNPFISFGLPKOPAWGITSALDKTMEIYWKRIIGTPHRS 694
 444 -----PV-----LSTGGVLFNGGT--DRYFRALSEQETLMQTRLATVVS 483
 695 MPFMPVVPFPMFGMPLGCGISTAGNVLFIAATADYLVVVISAGGHSFGT--AG 752
 484 GQA--ISYEVDGMOYVAI--AGGVSYS 508
 753 GQATPMFTEVNGKQYVVISAGGHSFGT 780

RESULT 13

S00943
 glucose dehydrogenase (pyrroloquinoline-quinone) (EC 1.1.99.17) - Acinetobacter calcoace
 C:Species: Acinetobacter calcoacetiscus
 C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
 C:Accession: S00943
 R:Cletoen-Jansen, A.M.; Goosen, N.; Odle, G.; van de Putte, P.
 Nucleic Acids Res. 16, 6228, 1988
 A:Title: Nucleotide sequence of the gene coding for quinoprotein glucose dehydrogenase
 A:Reference number: S00943; MUID:8829368
 A:Accession: S00943
 A:Molecule type: DNA
 A:Residues: 1-801 <CLE>
 A:Cross-references: EMBL:X07235; NID:g38711; PID:CAA50222.1; PID:g38712
 A:Experimental source: strain LMD 79.41
 C:Function:
 A:Description: catalyzes the oxidation of D-glucose to D-gluconic acid by ubiquinone
 C:Pathway: respiratory chain
 C:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)
 C:Keywords: oxidoreductase; pyrroloquinoline quinone; respiratory chain; transmembrane
 F:9-35/Domain: transmembrane #status predicted <TM1>
 F:39-57/Domain: transmembrane #status predicted <TM2>
 F:61-79/Domain: transmembrane #status predicted <TM3>
 F:94-108/Domain: transmembrane #status predicted <TM4>
 F:118-137/Domain: transmembrane #status predicted <TM5>
 F:91,93/Binding site: ubiquinone (Arg, Asp) #status predicted
 F:471/Active site: Asp #status predicted

Query Match 12.0%; Score 329; DB 1; Length 801;
 Best Local Similarity 21.9%; Pred. No. 1e-15;
 Matches 152; Conservative 81; Mismatches 197; Indels 264; Gaps 33;

20 PAFQVPTVDELLANPAGWISYQONENRHSPLQITENYGOLOVWARMQMGK 79
 152 PETAQAVPVAAE-----SDMPAGRGQAGRYSPLOINDONVKLAWT--LRGD 202
 80 V-----QVPLIHGQVWYLANPEDIQAIADANTGDLW-----EH-- 114
 203 LKTDNDGSETTQVPIKIGNMMFICTAHQOLAIIDPATGKEKMFDPKLTGDSFQHLT 262
 115 -----RROL-----NIATLNS-----FEE-- 129
 263 CGWVYDANNTTERATSIQSKSSSTCCPRKRVFVNDGRLVAVNDGACDFFQNG 322
 130 -----PTRGMAIVANGVIVAGS--TCQYS--PFQGVSGHDSATGEEL 168
 323 QVNLQEMFYAPYAGYNTSPGIYGVSTVYAGSVTDVNSKKESSG--VIRGIDVNTCKL 381
 169 WNNYFIPRAGE-----GGDETGWGNDYARMMTGAKGQITDPVTNLVHYGSTAVGPASE 222
 382 W--VEDTGADPNAMPGEKSTFVHNS-----PNAMAPLAYDAKLIV--YVPTGV----- 427
 223 TQGTGGTLYG-----TMTFAVPRDTEGIWVRHQTLPBWDWDECEFFEMAVTN 272
 428 ---GTP--DINGGDRTELKERYANSMLAINASTGKLVNFOTHHDLMDVPSQSLAD 482
 273 VDVPSTMEBLOSINP--NAATGERVLTGVP-----CITG--TMMQFPAET 316
 483 IKNRAGOTVPAIYVLTGNAFVLDNR--NGQPLVPTKRPQYTKRGGQTGGEYSKT 540
 317 GEF-----LMA-----RDTNINOMTESIDENGIVTVNDADIL 348
 541 QPESDLNLAPODKLTQDKMGATMLDQIMCRVSFKRLNLYGIVTPPSENGTL----- 592
 349 KELDVEYDCEPTFLGRDPSAALNPS-----GIYVI----- 381
 593 ---VPPGNLGVFEMGKMSVNDPROVAVNPIGLPFLVSLIPADPNRAQTAKAGCT 644
 382 ---PLANNCYDMAVDEFTSMADYNTSNVTKLPKGMIDGIRDAIDISTGRTLM-- 433
 645 EGVQVPMYGVY-----GVEISAFLSPLGLPCQKRPANGYVAGVDLKTHEVWKKR 694
 434 -----SVERAANYSPLVLTGGGVLENGST--DRYFRALSOETGETIMQTR 477

Db 695 IGTIRDSPLNLFQPAVKIGVPGIGSISIRAGVNMVFGATQDVLAFVNTGKRLWEAR 754
 QY 478 LATVASGA--ISIEVGMQYVAI--AGGVSYGS 508
 Db 755 LP--AGQADPMYEINGKQYVIMAGHGSGFCT 786

RESULT 14 OPERX

glucose dehydrogenase (pyrroloquinoline-quinone) (EC 1.1.99.17) - gluconobacter oxyda
 C:Species: Gluconobacter oxydans
 C>Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 22-May-1998
 C:Accession: S17716; S19265
 R:Cletoen-Jansen, A.M.; Dekker, S.; van de Putte, P.; Goosen, N.
 Mol. Gen. Genet. 229, 206-212, 1991
 A:Title: A single amino acid substitution changes the substrate specificity of quinop
 A:Reference number: S17716; MUID:92017653
 A:Accession: S17716
 A:Molecule type: DNA
 A:Residues: 1-808 <CLE>
 A:Cross-references: EMBL:X62710
 R:Goosen, N.
 submitted to the EMBL Data Library, February 1992
 A:Reference number: S19265
 A:Accession: S19265
 A:Molecule type: DNA
 A:Residues: 1-212, 'A', 214-808 <GOO>
 A:Cross-references: EMBL:X62710; NID:g58416; PID:g58417
 C:Genetics:
 A:Gene: gdh
 C:Function:
 A:Description: catalyzes the oxidation of D-glucose to D-gluconic acid by ubiquinone
 A:Pathway: respiratory chain
 C:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)
 C:Keywords: oxidoreductase; pyrroloquinoline quinone; respiratory chain; transmembran
 F:9-28/Domain: transmembrane #status predicted <TM1>
 F:35-54/Domain: transmembrane #status predicted <TM2>
 F:60-76/Domain: transmembrane #status predicted <TM3>
 F:94-110/Domain: transmembrane #status predicted <TM4>
 F:122-138/Domain: transmembrane #status predicted <TM5>
 F:91,93/Binding site: ubiquinone (Arg, Asp) #status predicted
 F:470/Active site: Asp #status predicted

Query Match 10.7%; Score 294.5; DB 1; Length 808;
 Best Local Similarity 20.7%; Pred. No. 3.1e-13;
 Matches 145; Conservative 82; Mismatches 234; Indels 241; Gaps 27;

13 ALALLAFAFQVPTVDELLA-----NPRAGWISYQONENRHSPLQITTT 61
 128 AVLAFLASLFTPDHDSIGELPTQIATNASPADPDVNPASEHNAVGRQAGRMSPLNOINA 187
 62 ENVGQLOLVW-----ARGMGPKV--QVPLIHGQVWYLANPEDIQAIADANTGDLW 112
 188 TIVNSLTKAMHIIHKDMANSNDPEQNTENKPTLEFNNLTVMCSLHOKLRAVDGAGGVNM 247
 113 EHRQOL--PNIAATLN--SFGE-PTRCM-----AIYANGVIV----- 143
 248 VYDPLQINPFOHLTGCGVSPHETPANMDSQNPAPDCAKDSITLPIVNDGRILVEDAD 307
 144 AGSTC-----QYSPGCEV----- 157
 308 TGKTCSEFGNGGEIDLVRPNQPTTPGQYEPTSPVITDKLIANSATINDGNSVQASGA 367
 158 -SGHDSATGEELW-----RNFFIPRAGEEDDETGMNDYERMMTGAKGAGQITTPVNTLVHY 212
 368 TQAFDYVTGKRWVVDASNDPQDLDESHPVFHPSPNSMTVS-----SYDANILVYI 422
 213 GSTAVGPASETQGTGTP--GGTLYGNTFRF-----AVRPDTEGIWVRHQTLPBWDWDECE 264
 423 PMGV-----GTPDQNGGDRTKDSEKFAFGIYALNADTKLAMPYQYVHHDLMDLME 473

```

QY 265 TREMATVNDVDPSTEMEGLOSINPNAATGERRVLITGPKCTGTMQDFAETG-EFLMAR 323
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 474 PSQPSIVDTOKDGLVPAIYA-----PRTGDI FVLDRTREKEIVPAP 517
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 324 DT-----NYQNMIESIDENGIVTVNEDALIKELDY-----EYD 356
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 518 EFPVPOGAAPGCHTSTQPMQSQ---LTLRPKNPLNDSDIMGCTIFDQMFCSYFHTLRYE 574
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 357 -----VCPFLGGRDMPSAALNPDGIFYIFPLNNVCYDMAAVDO----- 395
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 575 GPETPPLSGSLFPGDLQMFEMGGLAVDPQROVAFANPDISLPEVSQLVPRGPNPLMPE 634
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 396 -----EFTSMDVYNTSNVTKLPKGMKMI-----GRIDAIDISTGRTL 432
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 635 ENAKGTGETGLQHNHGIPIYAVNLHPFLDPVLLPFGIMPKCTPPWGVAGIDELKTNKV 694
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 433 MSVERAANYS-----PV-----LSTGGVLEF-NGGTDYFRALSOETG 470
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 695 WQHRNGLTNDMSWSSLPPLPPKIGVPSLGLPLSTAGNLGFLTASMDYIIRAVNLTYG 754
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 471 ETLMTRLATVASGQALSTVDGMQYVAIAGGVSYGSLNS 512
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 755 KVLMDRLPAGAOATPIYAINKQYI-----VTYAGSHNS 790
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

RESULT 15

```

F83360
glucose dehydrogenase PA2290 [imported] - Pseudomonas aeruginosa (strain PA01)
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: F83360
R:Stover, C.K.; Pham, X.O.; Ewlin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Ba-
  adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Llim,
  ; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho-
A:Reference number: AB2950; M01D:20437337
A:Accession: F83360
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-803 <STO>
A:Cross-references: GB:AE004654; GB:AE004091; MID:g9948311; PIDN:AA05678.1; GSPDB:GN001
A:Experimental source: strain PA01
C:Genetics:
A:Gene: gcd
C:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)

```

Query Match 10.28; Score 279; DB 2; Length 803;

Best Local Similarity 20.7%; Pred. No. 4.1e-12; Mismatches 226; Indels 276; Gaps 32;

Matches 152; Conservative 79;

```

QY 6 LLM-----ASAGALALLAPAFPAQVPTVDLLANP----- 36
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 103 LLMLEFRRPLADGPAPLGTALCAVAVLAGAAVGSQ-FTNPGLIVGRIDRDSGMTST 161
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 37 ----PAGEMISYQONQENYRHSPLQITTEENVGQLQVWARGMQPKV-----Q 81
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 162 APAMDGDMQAYGRTEFGDRYSPLKQITPANVQGLEAWR--IRTGDLPTADDPLELTINE 219
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 82 VTPLIHDGVMTLANPDVITQALDAKTGDLTWEHRQLPNATLNSFGEPT-RGM----- 134
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 220 NTPLVNGLVACTAHSKVTALADPDTGAETWRFDQIOSPVGEKGFAMTCRGVSYYDEE 279
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 135 -----AIVANGVIVA-----GSTCQ----- 149
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 280 QYARSDVCAFPALAEAGKAAVAASCPRLFLPTADARLAINADNGKVCEDFGKAVDL 339
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 150 -----YSDFGCF-----VSQH---DSATGEE--LMKNYFIPRAGEGDETW 185
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 340 TAGIGFTPGGYTSPAAVTRNLVLIIGHYTDNESTNEPSGVIRAFDY---HDGKLW 395
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 186 GND-----YEARWMTGAMGQIITYDPVTNLVHYGSTAVGPASETQRG---TP 228
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

Db 396 NMDSGNPDETEPLADGKFTYTRNSPNMMSLASVDEKLGQV---LPLDNQMPDQNGKRT 452
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 229 GGTILGNTNRFVVRPDGTGEIYWRHQTLPDNNMDQCFEMAVTVNDVDPSTEMEGLOSIN 288
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 453 GAERKSAGI-VALDINTGKLRNNYQFTTHLDMD-----VGSQPTL-----LD 495
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 289 PNAATGERRVLTVGPKCTGTMQDFAETG----- 317
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 496 LKTAQGVKALL-APTKQGSILVLDLRDGTPIVPIREVPAPQGAVEGDHPTAPQARSDLN 554
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 318 -----EFLMA-----RQTNQNMIESIDENGIVTVNEDALIKELDYEYD 356
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 555 LNRPLTERDMMGSSPEFDMCLRIQFRSLRYBGQYTPSBOGSL----- 598
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 357 VCPFLGGRDMPSAALNPDGIFYIFPLNNVCYDMAAVDOEFT---SMDVYNTSNVK--- 410
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 599 IYGNQNGVFNWGVSVDPVROLFTSPNMAFVSQMPRDKVPSGSKREGETSGVQPTG 658
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 411 -----LPPKDMIGRIDAIDISTGRTLSVERAAA-NYSPV----- 445
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 659 APYAVIMHPFMSPIGLPCQAPSWGDVAGIDLTITAKVWQHKNGTISRDMTPVPIGLTVGP 718
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 446 ----LSTGGVLEFNGGT-DRYFRALSOETGETLMTQRLATVASGOA--ISYE-VDGMQY 496
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 719 SMGGSITTAGVAFILSGTLDQYIRAYDAVDKQIMQARLP--AGQATPMYSYTGKDGROY 776
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 497 VAI-AGGGSYGS 508
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 777 VLIYAGHGSFGT 789
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

Search completed: May 24, 2002, 10:21:10
Job time: 348 sec

Fri May 24 11:27:33 2002

wallick-934-135.pep.rpr

Page 10

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2002, 10:30:54 ; Search time 28.06 Seconds

(Without alignments) updates/sec
706.501 Million cell

WALICK-934-135.PEP

Title: 2742
Perfect score: 1 MKPSSLIMASGALALLAAP.....GMOYVAAGGVSYSGSLNS 512

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	488	17.8	739	1 DHET_ACEBU	044002 acetobacter
2	483.5	17.6	738	1 DHET_ACEPO	P28936 acetobacter
3	483	17.6	742	1 DHET_ACEAC	P18278 acetobacter
4	473.5	17.3	757	1 DHET_GLUOX	005542 gluconobact
5	441.5	16.1	600	1 KOXE_PARDE	P29968 paracoccus
6	434.5	15.8	626	1 DHM1_METEX	P16027 methyllobact
7	433	15.8	631	1 DHM1_PARDE	P12293 paracoccus
8	430.5	15.7	626	1 DHM1_METOR	P15279 methyllobact
9	417	15.2	623	1 EXAA_PSEAE	Q924J7 pseudomonas
10	389.5	14.2	571	1 DHM1_METME	P38539 methylolophil
11	349.5	12.7	766	1 DHG_ECOLI	P15877 escherichia
12	329	12.0	801	1 DHGA_ACTICA	P05465 acetobact
13	294.5	10.7	808	1 DHG_GLUOX	P27175 gluconobact
14	254.5	9.3	809	1 QUILA_ACTICA	059086 acetobact
15	235	8.6	790	1 QUILA_XANCU	Q9X678 xanthomonas
16	136.5	5.0	392	1 YFGL_ECOLI	P77774 escherichia
17	125.5	4.6	827	1 AFSK_STRCO	Q6082 methanococc
18	118.5	4.3	799	1 AFSK_STRCO	Q6082 methanococc
19	116	4.2	3567	1 FLEY_CAUCR	003132 saccharopol
20	115.5	4.2	353	1 YXAL_BACSU	P42111 bacillus su
21	115.5	4.2	954	1 FLEY_CAUCR	P15335 cauliobacter
22	115.5	4.2	3354	1 CADN_MOUSE	Q996T4 mus musculu
23	115	4.2	443	1 PORD_PSEAE	P32722 pseudomonas
24	114	4.2	796	1 COPP_SCHPO	042937 schizosacch
25	113	4.1	621	1 ASPA_AERSA	031339 aeromonas s
26	112.5	4.1	1355	1 PPOD_AANSP	P22705 anabaena sp
27	112	4.1	790	1 PLAM_PIG	P06867 sus scrofa
28	111	4.0	415	1 Y232_RICPR	Q924J2 rickettsia
29	111	4.0	1637	1 MRSP_STAU	P80544 staphylococ
30	111	4.0	1645	1 MRSP_RICTY	P96989 r outer mem
31	110.5	4.0	593	1 SPG2_STRSP	P19909 streptococ
32	110.5	4.0	3317	1 CADN_RAT	P58365 rattus norv
33	110	4.0	649	1 ACES_DROME	P07140 drosophila

34	108.5	4.0	3354	1 CADN_HUMAN	O9h251 homo sapien
35	107.5	3.9	595	1 YB85_SCHPO	O14301 schizosacch
36	106.5	3.9	466	1 MM08_RAT	O88766 rattus norv
37	105	3.8	465	1 MM08_MOUSE	O70138 mus musculu
38	105	3.8	1365	1 GFES_STRDO	P29336 streptococ
39	104.5	3.8	827	1 XANP_XAN2	O60106 xanthomonas
40	102	3.7	697	1 NAME_STRPN	O54727 streptococ
41	101.5	3.7	1256	1 FINE_CHICK	P11722 gallus gall
42	101	3.7	1012	1 POLG_IBDO	P72726 avian infec
43	100	3.6	566	1 AMY_STRGR	P30270 streptomyce
44	100	3.6	985	1 INVA_YERPS	P11922 yersinia ps
45	99.5	3.6	677	1 SNAC_BACSU	P05656 bacillus su

ALIGNMENTS

RESULT 1	ID	DHET_ACEBU	STANDARD:	PRT:	739 AA.
AC	Q44002	007952:			
AC	15-DEC-1998	(Rel. 37, Created)			
DT	15-DEC-1998	(Rel. 37, last sequence update)			
DT	15-JUL-1999	(Rel. 38, last annotation update)			
DE	Alcohol dehydrogenase [acceptor] precursor (EC 1.1.99.8).				
GN	ADH.				
OS	Acetobacter europaeus.				
OC	Bacteria; Proteobacteria; alpha subdivision; Acetobacteraceae;				
OC	Gluconobacter.				
OX	NCBI_TaxID=33995;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=DES11 / DSM 6160;				
RA	Thurner C.A.K.;				
RL	Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.				
CC	-1- CATALYTIC ACTIVITY: A primary alcohol + acceptor = an aldehyde + reduced acceptor.				
CC	-1- COFACTOR: POQ AND HEME (BY SIMILARITY).				
CC	-1- SUBUNIT: HETEROTETRAMER (DEHYDROGENASE, CYTOCHROME, AND TWO SMALLER UNKNOWN SUBUNITS) THAT FORMS THE ALCOHOL DEHYDROGENASE COMPLEX (BY SIMILARITY).				
CC	-1- SUBCELLULAR LOCATION: MEMBRANE-BOUND, FACING THE PERIPLASMIC SPACE (POTENTIAL).				
CC	-1- SIMILARITY: BELONGS TO THE BACTERIAL POQ DEHYDROGENASE FAMILY.				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).				
CC	-----				
DR	EMBL; X82894; CA58066.1; -				
DR	EMBL; Y09480; CA70688.1; -				
DR	HSSP; Q924J7; IFLG.				
DR	InterPro; IPR001479; Bac_POQ.				
DR	InterPro; IPR002372; Bac_POQ_repeat.				
DR	InterPro; IPR000345; CytC_heme_bind.				
DR	Pfam; PF01011; Bacterial_POQ_6.				
DR	PROSITE; PS00363; BACTERIAL_POQ_1; 1.				
DR	PROSITE; PS00364; BACTERIAL_POQ_2; 1.				
DR	PROSITE; PS00190; CYTOCHROME_C; 1.				
KW	Oxidoreductase; POQ; Heme; Periplasmic; Membrane; Signal.				
FT	SIGNAL	1	35	POTENTIAL.	
FT	CHAIN	36	739	ALCOHOL DEHYDROGENASE [ACCEPTOR].	
FT	BINDING	651	651	HEME (COVALENT) (BY SIMILARITY).	
FT	BINDING	654	654	HEME (COVALENT) (BY SIMILARITY).	
FT	METAL	655	655	IRON (HEME AXIAL LIGAND) (BY SIMILARITY).	
SQ	SEQUENCE	739 AA:	80944 MM:	EG81BB237ACB91F4 CRC64;	

Query Match

17.8%, Score 488; DB 1; Length 739;

Best Local Similarity 26.0%; Pred. No. 8.9e-29;
Matches 152; Conservativeness 76; Mismatches 246; Indels 110; Gaps 14;

```

OY 10 SAGALALLAAPAFQVPTVDE-----LLANPAGCEWISYQONENYHSLPQITTT 61
DB 17 TAGTICALLISGYATMASADGOGATGEAIIHADDPHNNMTYGRYSEGRYSPLDQINR 76
OY 62 ENVGOLQIWMARGOPGVQYV-TPLIHDGVYVLANPGDVIAIDAKTGDIIMHRRQLP- 119
DB 77 SNVGNLKLAWYLDLDTNRGOGCTPLVIDGVYATNTNMSMKAVDAATGKLMSYDPRVPG 136
OY 120 NATLNSGEPTRGMAI----- 136
DB 137 NIADKCCDVTYVNGAAYWNGKVFEGTEPDRLLALDAKTGLVMSVNTIPPEALGKORSY 196
OY 137 -----VANGVIAGSTCOYSPGCC--FVSGHDSATGEELMANTYTPRAGEBDETWGN 187
DB 197 TYDGAAPRIAKGRVILGN--GGSEFGARGFYATDAETGKVDWRFETAPRKNPDHTASD 254
OY 188 D-----YEARWMTGA-----WGOITYPVTNLVHYGSTAVGPASETQRTPGGT 231
DB 255 SYLANKAYQWTSPTGANTRGGGGTWDSIVYDPAVDLYLVGNGSPNNYRSEBKGCD 314
OY 232 LYGNTEFAVAPDTGELVYHQTLPNDNDQECTFEEMVTNVDPSTMEGLQISINPA 291
DB 315 NLFTGSIVALKPTEGTYVWHFOETPMQDFTSVQOIMTLDPINGET----- 362
OY 292 ATGERRVLTGVPCKGTGMMOFDAETGEPLMARDNTYONNIESID-ENGIVTYNEDAILKE 350
DB 363 -----RHYIVHAP-KNGEFTIIDAKTGEFTISGKNYVYVWNASGLDPTGPIYVNPALYTL 417
OY 351 LDVEYDVCTPFLGDRWPSALNPDGSIYFPLNNVCYDMAVDOEFT-SMDVNTS--- 406
DB 418 TGEWYGPIDGLGHNFAAFESPRTGLVYIPAOQVPLTYNQVGGTFPHRPSMWLGIDM 477
OY 407 NVTKLRPG-----KDMIGRIDALIDISTGRTLSVEPAANISPVLSGGVLENGGT 458
DB 478 NKAGIDSPBAKQAFVKDLKGMVAMPDQKQAEAMRVHKKFPWNGGILATGDLFLQCLA 537
OY 459 DRYFRALSOETGETLQWTRLATYVAGSAISYEVDGMQVYAIAGG 502
DB 538 NGEFHAYDATNGSDLFHFAADSGIAPRYTYLANKQYVAVEVG 581

RESULT 2
DHET ACEPO STANDARD: PRT: 738 AA.
AC P28036:
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Alcohol dehydrogenase [acceptor] precursor (EC 1.1.99.8).
GN ADH.
OS Acetobacter polyoxogenes.
OC Bacteria; Proteobacteria; alpha subdivision; Acetobacteraceae;
OC Acetobacter.
OX NCBI_TaxID=439;
RN [1]
RP SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.
RC STRAIN=NB11028;
RX MEDLINE=91159482; PubMed=2001402;
RA Tanaka T., Fukaya M., Takemura H., Tayama K., Okumura H.,
RA Kawamura Y., Nishiyama M., Horinouchi S., Beppu T.;
RT "Cloning and sequencing of the gene cluster encoding two subunits of
RT membrane-bound alcohol dehydrogenase from Acetobacter polyoxogenes.";
RL Biochim. Biophys. Acta 1088:292-300(1991).
CC -1- FUNCTION: DEHYDRATION OF PRIMARY ALCOHOLS (EXCEPTION: METHANOL).
CC -1- CATALYTIC ACTIVITY: A primary alcohol + acceptor = an aldehyde +
CC reduced acceptor.
CC -1- COFACTOR: POO AND HEME.
CC -1- SUBUNIT: HETEROTETRAMER (DEHYDROGENASE, CYTOCHROME, AND TWO
CC SMALLER UNKNOWN SUBUNITS) THAT FORMS THE ALCOHOL DEHYDROGENASE
CC COMPLEX.

```

-1- SUBCELLULAR LOCATION: MEMBRANE-BOUND, FACING THE PERIPLASMIC SPACE (POTENTIAL).
-1- SIMILARITY: BELONGS TO THE BACTERIAL POO DEHYDROGENASE FAMILY.
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

```

CC EMBL: D00635; BAA0528.1; -
CC PIR: Q14270; S14270.
CC HSP: Q924J7; 1P1G.
CC InterPro: IPR001479; Bac_POO.
CC InterPro: IPR002372; Bac_POO.
CC InterPro: IPR000345; CytC_heme_bind.
CC Pfam: PF01011; Bacterial_POO_6.
CC PROSITE: PS00363; BACTERIAL_POO_1; 1.
CC PROSITE: PS00364; BACTERIAL_POO_2; 1.
CC PROSITE: PS00190; CYTOCHROME_C_1.
CC Oxidoreductase; POO; Heme; Periplasmic; Membrane; Signal.
CC FT SIGNAL 1 35
CC CHAIN 36 738
CC BINDING 650 738 ALCOHOL DEHYDROGENASE [ACCEPTOR].
CC BINDING 650 650 HEME (COVALENT) (BY SIMILARITY).
CC METAL 654 654 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
CC SEQUENCE 738 AA; 80840 MW; 1E2B6ED7BCD92A66 CMC64;

```

Query Match 17.6%; Score 483.5; DB 1; Length 738;
Best Local Similarity 25.8%; Pred. No. 1.9e-28;
Matches 152; Conservativeness 76; Mismatches 239; Indels 123; Gaps 14;

```

OY 10 SAGALALLAAPAFQVPTVDE-----LLANPAGCEWISYQONENYHSLPQITTT 61
DB 17 TAGTICALLISGYATMASADGOGATGEAIIHADDPHNNMTYGRYSDORYSPLDQINR 76
OY 62 ENVGOLQIWMARGOPGVQYV-TPLIHDGVYVLANPGDVIAIDAKTGDIIMHRRQLP- 119
DB 77 SNVGNLKLAWYLDLDTNRGOGCTPLVIDGVYATNTNMSMKAVDAATGKLMSYDPRVPG 136
OY 120 NATLNSGEPTRGMAI----- 136
DB 137 NIADKCCDVTYVNGAAYWNGKVFEGTEPDRLLALDAKTGLVMSVNTIPPEALGKORSY 196
OY 137 -----VANGVIAGSTCOYSPGCC--FVSGHDSATGEELMANTYTPRAGEBDETWGN 187
DB 197 TYDGAAPRIAKGRVILGN--GGSEFGARGFYATDAETGKVDWRFETAPRKNPDHTASD 254
OY 183 -----YEARWMTGA-----WGOITYPVTNLVHYGSTAVGPASETQRTPGGT 231
DB 255 SYLANKAYQWTSPTGANTRGGGGTWDSIVYDPAVDLYLVGNGSPNNYRSEBKGCD 314
OY 232 LYGNTEFAVAPDTGELVYHQTLPNDNDQECTFEEMVTNVDPSTMEGLQISINPA 291
DB 315 NLFTGSIVALKPTEGTYVWHFOETPMQDFTSVQOIMTLDPINGET----- 362
OY 292 ATGERRVLTGVPCKGTGMMOFDAETGEPLMARDNTYONNIESID-ENGIVTYNEDAILKE 350
DB 363 -----RHYIVHAP-KNGEFTIIDAKTGEFTISGKNYVYVWNASGLDPTGPIYVNPALYTL 417
OY 351 LDVEYDVCTPFLGDRWPSALNPDGSIYFPLNNVCYDMAVDOEFT-SMDVNTS--- 406
DB 418 TGEWYGPIDGLGHNFAAFESPRTGLVYIPAOQVPLTYNQVGGTFPHRPSMWLGIDM 477
OY 407 NVTKLRPG-----KDMIGRIDALIDISTGRTLSVEPAANISPVLSGGVLENGGT 458
DB 478 NKAGIDSPBAKQAFVKDLKGMVAMPDQKQAEAMRVHKKFPWNGGILATGDLFLQCLA 537
OY 459 DRYFRALSOETGETLQWTRLATYVAGSAISYEVDGMQVYAIAGG 502
DB 538 NGEFHAYDATNGSDLFHFAADSGIAPRYTYLANKQYVAVEVG 581

RESULT 2
DHET ACEPO STANDARD: PRT: 738 AA.
AC P28036:
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Alcohol dehydrogenase [acceptor] precursor (EC 1.1.99.8).
GN ADH.
OS Acetobacter polyoxogenes.
OC Bacteria; Proteobacteria; alpha subdivision; Acetobacteraceae;
OC Acetobacter.
OX NCBI_TaxID=439;
RN [1]
RP SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.
RC STRAIN=NB11028;
RX MEDLINE=91159482; PubMed=2001402;
RA Tanaka T., Fukaya M., Takemura H., Tayama K., Okumura H.,
RA Kawamura Y., Nishiyama M., Horinouchi S., Beppu T.;
RT "Cloning and sequencing of the gene cluster encoding two subunits of
RT membrane-bound alcohol dehydrogenase from Acetobacter polyoxogenes.";
RL Biochim. Biophys. Acta 1088:292-300(1991).
CC -1- FUNCTION: DEHYDRATION OF PRIMARY ALCOHOLS (EXCEPTION: METHANOL).
CC -1- CATALYTIC ACTIVITY: A primary alcohol + acceptor = an aldehyde +
CC reduced acceptor.
CC -1- COFACTOR: POO AND HEME.
CC -1- SUBUNIT: HETEROTETRAMER (DEHYDROGENASE, CYTOCHROME, AND TWO
CC SMALLER UNKNOWN SUBUNITS) THAT FORMS THE ALCOHOL DEHYDROGENASE
CC COMPLEX.

```

DB 531 LFOGLANGFHAHYDATINGSDLEFHAADSGIAPPTYLANGKQYVAEVC 580

RESULT 3

DHET_ACAC STANDARD: PRT; 742 AA.

AC 18278;

DT 01-NOV-1990 (Rel. 16, Created)

DT 01-NOV-1990 (Rel. 16, Last sequence update)

DT 15-DEC-1998 (Rel. 37, Last annotation update)

DE Alcohol dehydrogenase [acceptor] precursor (EC 1.1.99.8).

GN ADHA OR ADH.

OC Acetobacter acet.

OC Bacteria; Proteobacteria; alpha subdivision; Acetobacteraceae;

OC Acetobacter.

OX NCBI_Taxid=435;

[1]

RP SEQUENCE FROM N.A., AND SEQUENCE OF 36-50.

RX MEDLINE=6925070; PubMed=2722742;

RA Inoue T., Sunagawa M., Mori A., Imai C., Fukuda M., Takagi M., Yano K.;

RT "Cloning and sequencing of the gene encoding the 72-kilodalton dehydrogenase subunit of alcohol dehydrogenase from Acetobacter acet.;"

RT Biochem. J. 308:375-379(1995).

RL Blochem. J. 308:375-379(1995).

[2]

RN J. Bacteriol. 171:3115-3122(1989).

RP 3D-STRUCTURE MODELING.

RX MEDLINE=9528964; PubMed=7772016;

RA Collier G.E., Gilles I.G., Anthony C.;

RT "The structure of the quinoprotein alcohol dehydrogenase of Acetobacter acet. modelled on that of methanol dehydrogenase from Methylobacterium extorquens.;"

RL Biochem. J. 308:375-379(1995).

CC -1- CATALYTIC ACTIVITY: A primary alcohol + acceptor - an aldehyde + reduced acceptor.

CC -1- COFACTOR: POQ AND HEME.

CC -1- SUBUNIT: TRIMER OF NON IDENTICAL CHAINS (DEHYDROGENASE, CYTOCHROME, AND TWO SMALLER UNKNOWN SUBUNITS).

CC -1- SUBCELLULAR LOCATION: Periplasmic.

CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POQ DEHYDROGENASE FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL: D90004; BAA14058.1; -

DR PIR: J50326; J50326.

DR HSSP: Q9Z4U7; 1FLG.

DR InterPro: IPR001479; Bac_POQ.

DR InterPro: IPR002372; Bac_POQ_repeat.

DR InterPro: IPR000345; CytC_heme_bind.

DR Pfam: PF01011; Bacterial_POQ_6.

DR PROSITE: PS00363; BACTERIAL_POQ_1; 1.

DR PROSITE: PS00364; BACTERIAL_POQ_2; 1.

DR PROSITE: PS00190; CYTOCHROME_C; 1.

DR OXidoreductase; POQ; Heme; Periplasmic; Membrane; Signal.

KW SIGNAL

FT CHAIN 1 35

FT ACT_SITE 343 742 ALCOHOL DEHYDROGENASE [ACCEPTOR].

FT BINDING 649 649 BASE (POTENTIAL).

FT BINDING 652 652 HEME (COVALENT) (BY SIMILARITY).

FT METAL 653 653 HEME (COVALENT) (BY SIMILARITY).

FT METAL 653 653 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).

SEQUENCE 742 AA: 81521 MW; 96692660DAB825A CRC64;

Query Match 17.6%; Score 483; DB 1; Length 742;

Best Local Similarity 26.5%; Pred. No. 2,1e-28;

Matches 163; Conservative 88; Mismatches 208; Indels 156; Gaps 24;

QY 11 AGALALAPAFAGVPTVTDLEL--ANPPAGEWISYQONDENRHSPLTQITTEENGOLO 68

DB 23 AALALPAAVAPADGQGNTEGAILHADHPENWLSYGRYSQRYSPLDQINRSNVGDLK 82

QY 69 LVWARGMOPGKVQ-VPTLIHGVMTLANPGVYQAIDAKTGDLMEHRROLP-NIA----- 122

DB 83 LIGYTYLDTNRQGEATPLVVDGIMVATTNWSKMEALDAAATKLIQIDPKVGNIAQKC 142

QY 123 --TLN-----SEGE-----PTRGMAI 136

DB 143 CDTVNRGAGTWNMGKVPWFQFDPDRLYAADAATCKKQWAVNTIPADAISIGKQSYVDGAVR 202

QY 137 VANGVIVAGSTCOYSEFG--FVSGHDSATGEELRNPTIRAGEEGD----- 182

DB 203 VAKGLVLIGN--GGAFFGARGFVSAFDAETGKLRFPYTPNNKNEPDAASDNIANKA 260

QY 183 -ETWAGDYERKMT-----GANGQITYPDVTNLVYGSFVAVGPASETGPGTGLY--- 233

DB 261 YKTWGP--KGAWYRGCGGTWDSLVYDVSLLY---LAVG-----NGSPWNKRYSE 309

QY 234 --GTN-----TRFAVRDTEGIWRHQTLPDNDWDOCTEFEMVTVNVQPTSTEMGLQSI 287

DB 310 GIGSNLFLGSIYALKPTEGYVWHFQATPDMDQDYTSVQOIMTLDPYK----- 358

QY 288 NPNAAAGE-RRVLTGVPCKTGTMQFDATGETIARNDTNYQNMIESID-----EN 337

DB 359 -----GEMRHVIVHAP-KNGFEYVDAKTGFTSGKRVYQNMANGDPLTGRPNYMPD 411

QY 338 GIATVNVNDALIKELDEYDVCPFLGGRDMPSAALNPDGIFILPNVNCY----- 388

DB 412 GLYTLNG-----KRYTGI-PPGLAHNPMMAVSPKTHLVYIPAHQIPFGYKNQVGF 463

QY 389 ----DMAVDOFTSMQVNTVNTKLPPEKDMIGRIDALIDISTGTMTSYERRAANNSP 444

DB 464 KPHADSMNVGGLMTKNGLPDTPPE-ARTAYIKDLEHGLWAMDVPKMTVKIDHKGPWNG 522

QY 445 VLSTGGGVLRNGGDRFRFRLASQETGELMTQRLAVASQAISYEVDMQVYA----- 498

DB 523 ILATVGDLPLFOGLANGFHAHYDATINGSDLYKFFDAQSGIAPPTYVSNKQYVAEVC 582

QY 499 ----IAGGVSYGSG 509

DB 583 GIYIPMGVGVRTSG 597

RESULT 4

DHET_GLUOX STANDARD: PRT; 757 AA.

AC 005542;

DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Alcohol dehydrogenase [acceptor] precursor (EC 1.1.99.8) (G3-ADH subunit 1).

GN ADHA.

OC Gluconobacter oxydans (Gluconobacter suboxydans).

OC Bacteria; Proteobacteria; alpha subdivision; Acetobacteraceae;

OC Acetobacter.

OX NCBI_Taxid=442;

[1]

RP SEQUENCE FROM N.A., AND SEQUENCE OF 36-50.

RX STRAIN-IFO 12528;

RX MEDLINE=97208225; PubMed=9055427;

RA Kondo K., Horinouchi S.;

RT "Characterization of the genes encoding the three-component membrane-bound alcohol dehydrogenase from Gluconobacter suboxydans and their expression in Acetobacter pasteurianus.;"

RT Appl. Environ. Microbiol. 63:1131-1138(1997).

CC -1- FUNCTION: DEHYDRATION OF PRIMARY ALCOHOLS (EXCEPTION: METHANOL) (BY SIMILARITY).

CC -1- CATALYTIC ACTIVITY: A primary alcohol + acceptor = an aldehyde + reduced acceptor.

CC -1- COFACTOR: POQ AND HEME.


```

CC -1- SUBUNIT: HETEROTRIMER (DEHYDROGENASE, CYTOCHROME AND PROTEIN
CC ADHS), THAT FORMS THE ALCOHOL DEHYDROGENASE COMPLEX.
CC -1- SUBCELLULAR LOCATION: MEMBRANE-BOUND, FACING THE PERIPLASMIC
CC SPACE (POTENTIAL).
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POQ DEHYDROGENASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D86375; BAA19753.1; -.
DR HSSP; Q9Z4J7; 1F1G.
DR InterPro; IPR001479; Bac_POQ.
DR InterPro; IPR002372; Bac_POQ_repeat.
DR InterPro; IPR000345; Cyt_heme_bind.
DR Pfam; PF01011; Bacterial_POQ_6.
DR PROSITE; PS00363; BACTERIAL_POQ_1; 1.
DR PROSITE; PS00364; BACTERIAL_POQ_2; 1.
DR PROSITE; PS00190; CYTOCHROME_C; 1.
DR Oxidoreductase; POQ; Heme; Periplasmic; Membrane; Signal.
FT CHAIN 1 34
FT MOD_RES 35 757 ALCOHOL DEHYDROGENASE [ACCEPTOR].
FT ACT_SITE 342 342 PYRROLIDONE CARBOXYLIC ACID.
FT BINDING 653 653 HEME (POTENTIAL) (BY SIMILARITY).
FT BINDING 656 656 HEME (COVALENT) (BY SIMILARITY).
FT METAL 657 657 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
SQ SEQUENCE 757 AA; 82966 MW; 39B9F0E3B947581 CRC64;

Query Match 17.3%; Score 473.5; DB 1; Length 757;
Best Local Similarity 25.3%; Pred. No. 1.1e-27;
Matches 155; Conservative 88; Mismatches 233; Indels 137; Gaps 19;

OY 7 LMSAGALLIAA--PAFAOV---TPVDELLANPPAGEWISYSGONENYRHSPLDTQTT 61
DB 16 LLSAALAFSAVPAVPAFOEDPTGTAITSSDNGHP--DDWLSTYGRSYSEORHSPLDQJNT 74
OY 62 ENVGQLOLVWARGMOPGKQV--TPLIHDSVMYLANPGVIAIDAKTGDLWEHRROP- 119
DB 75 ENVGKILKAHMDLDITNRQEGTEPLIVNCKATITNWSKMAIDATGKLSMDPKVPG 134
OY 120 NATILNSEEPTRGMAI----- 136
DB 135 NIADRCCTDVSRAAAMNGKVFCTFDGRLIALDAKTGLWMSVYTIIPKFAQLGHQRSY 194
OY 137 -----VANGYIVAGSTQCYSPFC--FVSGHDSATGEELMRVFIIPRAGEED----- 182
DB 195 TVDGAAPRIAKGVILGN--GGAEFARFVSAFDETSKLDWREFTYVNPENKPKDGAASD 252
OY 183 -----ETWGNDEYEAAMWTG--AMQITTYDPTVNLVHSGTAVPASPSTQGRFGT 231
DB 253 DILMSKAVPTMGKAMKAMKGGGVTWDSLVYDPTDLVYLGANGSPMYKRRSEKGD 312
OY 232 LYGTFRAVVRDGEIYWRHQTLPKRWNOQECTFEKMTVINVDVOSTMEGLQSLNPA 291
DB 313 NLFGLSIVAINBDGKYVWHFOETPMDEMDTYSVOQINTLDMPY----- 356
OY 292 ATGE--RRVLTGVPCKTGTMTGMOFADTEGFEFLWARTNYQNMIESTE--NGIYTVNEADILK 349
DB 357 -NGEMRHVIYAP--KNGFFIIDAKTGKFTIGKRYTENNANGIDPVTGPNVVPDLMT 414
OY 350 ELADVEYDVCPTFLGGRDMPASALNPDGSIYFIPLNNV-----CYDMAVDQE 396
DB 415 LTGRKWLGPGLPEGLGHNFAAAYSPKTKLVYIPAOQIPLLYDGQKGFKAHYHAAWLGID 474
OY 397 FTSMDVYTSVNTKLPBGKMDIGRIDIDISTGRTL--WSEYRAAANS-----PVY 446
DB 475 MKKIGLFDNDNDEHVAAKDF-----LKVLKGMTVAMDEKMAPAFTINHKGPWNGILL 528

```

```

OY 447 STGGVLENGSTDRYFRALSOETGELMOTRLATVASGAISYVDGMQVAT----- 499
DB 529 ATAGNVIFQIANGEFAPHAVDATNNDLYSPFASAIAPVYITANGKQYAVENWGSI 588
OY 500 ---AGGVSYSYG 509
DB 569 YPFLXGVARTSG 601

RESULT 5
XOXF_PARDE STANDARD; PRT; 600 AA.
ID XOXF_PARDE
AC P29968;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Putative dehydrogenase XOXF precursor (EC 1.1.99.-).
GN XOXF.
OS Paracoccus denitrificans.
OC Bacteria, Proteobacteria, alpha subdivision, Rhodobacter group;
OC Paracoccus.
OX NCBI_TaxID=266;
RN [1]
RP SEQUENCE FROM N.A.
RA Harms N.;
RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
RP SEQUENCE OF 380-600 FROM N.A.
RC STRAIN=PD 1235;
RX MEDLINE=92041583; PubMed=1657873;
RA Ras J., Reijnders W.N.M., van Spanning R.J.M., Harms N., Oltmann L.F.,
RA Strouthamer A.H.;
RT "Isolation, sequencing, and mutagenesis of the gene encoding
RT cytochrome c531 of Paracoccus denitrificans and characterization of
RT the mutant strain."
RL J. Bacteriol. 173:6971-6979(1991).
CC -1- COFACTOR: POQ (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POQ DEHYDROGENASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U34346; AAC4455.1; -.
DR EMBL; M75583; AAA25574.1; -.
DR PIR; A41378; A41378.
DR HSSP; P38539; AAAH.
DR InterPro; IPR002372; Bac_POQ_repeat.
DR Pfam; PF01011; Bacterial_POQ_7.
KW Oxidoreductase; POQ; Signal.
FT SIGNAL 1 21
FT CHAIN 22 600 PUTATIVE DEHYDROGENASE XOXF.
FT ACT_SITE 318 318 BASE (POTENTIAL).
SQ SEQUENCE 600 AA; 65159 MW; DCA996F1BC5A3ACE CRC64;

Query Match 16.1%; Score 441.5; DB 1; Length 600;
Best Local Similarity 25.9%; Pred. No. 2e-25;
Matches 157; Conservative 84; Mismatches 220; Indels 145; Gaps 25;

OY 10 SAGALILAAFAFAQVTPVDELLANP-----PAGEWISYSGONENYRHSPLDTQITE 62
DB 6 NGACIALLMGSTAA-----LANQGRAGRDQAPQWAIQMGDYANRYSTLQDIND 56
OY 63 NVGQLOLVW--AROMQPGQVTPPLIHDSVMYLANP--GVYIOAIDAK--TGDILWEHR--RO 117
DB 57 NVKDLRAVAFPSFSGVLNGH--EGSPVIVGDMYVHTFPKRVFRAULDNDGKILMRYEPOO 115

```



```

DB 345 -----KDGARKLLTHPRNRIYVTLDRDGLVSAANKLDDT--VNVFKSVLDTGTQPR 397
QY 344 EDAILKELD-VEYDVCPTFLGRDMPASALNPDGCIYEFLNNVCYMAAVDDEFTSMGV 402
DB 398 DEYGTFRDHLAKDICPSAMGYHNOGHDSDPKRELFEMGINHICMDEPFMLPYRAGOF 457
QY 403 YNTSNVYTLPPCK-----DMGRIDAIDISTGRTLMSVRAAANYSPVLSTGGVLFNG 456
DB 458 FVGATLNNYPPKGDRONYEGIGQIKAYNAITGDYKMEKERRAFVWGCTATAGDLVYFG 517
QY 457 GTDRFRALSGEETLMTOTRIATVASSGAISEVDMGVYALAGGVSYGSG 509
DB 518 TLDTYKARSDTDGLMKRIFPSGALGIPYTHKGTQYVAI-----YGVG 565

RESULT 7
DHML_PARDE STANDARD; PRT; 631 AA.
AC P12293;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Methanol dehydrogenase subunit 1 precursor (EC 1.1.99.8) (MDH large
  alpha subunit) (MEDH).
GN MOXF.
OS Paracoccus denitrificans.
OC Bacteria; Proteobacteria; alpha subdivision; Rhodobacter group;
OC Paracoccus.
OX NCBI_TaxID=266;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 33-49.
RA MEDLINE-8707969; PubMed-3114231.
RX Harns N., de Vries G.E., Maurer K., Hoogendijk J., Scouthamer A.H.;
RT "Isolation and nucleotide sequence of the methanol dehydrogenase
  structural gene from Paracoccus denitrificans.";
RL J. Bacteriol. 169:3969-3975(1987).
CC -1- CATALYTIC ACTIVITY: A primary alcohol + acceptor = an aldehyde +
  reduced acceptor.
CC -1- COFACTOR: POO.
CC -1- SUBCELLULAR LOCATION: Periplasmic.
CC -1- MISCELLANEOUS: MDH IS THE MAJOR PROTEIN IN THE CELL DURING GROWTH
  ON METHANOL (IN P.DENITRIFICANS MDH CONSTITUTES UP TO 15% OF THE
  TOTAL CELL PROTEIN).
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POO DEHYDROGENASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M17339; AAA88366.1; -
DR HSSP: P38539; AAAH.
DR InterPro: IPR001479; Bac_POO.
DR InterPro: IPR002372; Bac_POO_repeat.
DR Pfam: PF01011; Bacterial_POO_7.
DR PROSITE: PS00363; BACTERIAL_POO_1; 1.
DR PROSITE: PS00364; BACTERIAL_POO_2; 1.
KW Oxidoreductase; POO; Signal; Methanol utilization; Periplasmic.
FT CHAIN 1 32
FT DISULFID 135 631 METHANOL DEHYDROGENASE SUBUNIT 1.
FT DISULFID 418 447 BY SIMILARITY.
FT ACT_SITE 335 335 BASE (POTENTIAL).
SEQUENCE 631 AA; 69799 MW; 0934DC93FFC5730B CMC64;

```

Query Match 15.8%; Score 433; DB 1; Length 631;
 Best Local Similarity 25.3%; Pred. No. 9,1e-25;
 Matches 152; Conservative 91; Mismatches 220; Indels 138; Gaps 25;

```

QY 4 TSLMASAGALALL-AAFAEQVTPVDEL--LANPAGEMISYGNONEYHRSPLTOT 60
DB 12 SLSMAVAVMGLAVLTTPATA-----NDQVELAKDPA-NWVTFGQDYNAQNTSEMTDIN 65
QY 61 TENNGQDLVAV-ARGMQPKGVOTPLIHGVWYLANP-GDVYQAIIDA-KTGDLIWEHR- 115
DB 66 KENYKOLRAVWSFSTGLVHGH-EGTPLYVDRMFHTFPNPTFALDNEPKIIMQNP 124
QY 116 ROLPNIATLNSFGEPTRGMA-----YANGVT 142
DB 125 KQNTARTVACCDVNNNGLAWPGDOVKPLIFRTOLDGIYAMALETETRMIMENDI 184
QY 143 VAGSTCOYSFP-----GC-----FVSGHDSATGEELMRYTF----- 174
DB 185 KVGSTLTIAPIYIKDLVLVSGSGAELGVGYVAYDVKSGEWRAPATGPDELLAED 244
QY 175 -----PRAGE--GDETWGNDYFARMGTG--KQITDPTNLVHGSTRVAGPASTQ 224
DB 245 FNAPNPHYGGKNGLETWEG--AKKIGGTMWYAYDPEVDLFYSGNPAPWNETM 301
QY 225 RGTPEGTLGTNTRFAVRDTEIWRHQTLPDNDQECTFEMWVTVDVQPTNEMGL 284
DB 302 R--PGDNKM-TMAIWGREATTGEAKFAVQKTPHDEM-----YAGVNMALSEQEDK 350
QY 285 OSINPMAATGERVVLGVPCCKTGMQPDATGEFLNARDNTYONMESIDE--NGIVTV 342
DB 351 Q-----GQNRKLLTHPRNRIYVTLDRNGDILISA-----DKMDPTVNWKEY 393
QY 343 NEDAILKELDVEY-----DVCPFLGGRDMPASALNPDGCIYFLPNNVCYDMAV 393
DB 394 QLDTGTPVADPEFRGMDHAKARDIPSAMGYHNOGHDSDYDERKVFMLGHNICMDEPF 453
QY 394 DEFTSMGVYNTYTKLPKPKDM-----IGRIDAIDISTGRTLMSVRAAANYSPVLST 448
DB 454 MLYRAGOFFGATLTMPGPRAATARAAGQAIKAYLAISGEMKEMEREFVWGTMAT 513
QY 449 GGGVLENGTDRYFRALSGEETLMTOTRIATVASSGAISEVDMGVYALAGGVSYGS 508
DB 514 AGCLTFEYTLDTGFIKARSDTDGLMKRIFPSGALGIPYTHKGTQYVAI-----MYGV 568
QY 509 G 509
DB 569 G 569

RESULT 8
DHML_METOR STANDARD; PRT; 626 AA.
AC P15279;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Methanol dehydrogenase subunit 1 precursor (EC 1.1.99.8) (MDH large
  alpha subunit) (MEDH).
GN MOXF.
OS Methylobacterium organophilum XX.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Methylobacterium group; Methylobacterium.
OX NCBI_TaxID=410;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 29-43.
RA STRAIN-ATCC 27886 / DSM 760 / NCIB 11278;
RX MEDLINE-89008094; PubMed-2459109;
RT "Nucleotide sequence and transcriptional start site of the
  Methylobacterium organophilum XX methanol dehydrogenase structural
  gene.";
RL J. Bacteriol. 170:4739-4747(1988).
CC -1- CATALYTIC ACTIVITY: A primary alcohol + acceptor = an aldehyde +
  reduced acceptor.
CC -1- COFACTOR: POO.
CC -1- SUBUNIT: THE HOLOENZYME MDH HAS AN ALPHA-2/BETA-2 CONFIGURATION
  OF A LARGE ALPHA AND A SMALL BETA SUBUNIT.

```

CC -1- SUBCELLULAR LOCATION: PERIPLASMIC, BUT ASSOCIATED WITH INNER
 CC MEMBRANE.
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POO DEHYDROGENASE FAMILY.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: M22629; AAA50289.1; -
 CC HSSP: P38539; AAAH.
 CC InterPro: IPR001479; Bac_POO.
 CC InterPro: IPR002372; Bac_POO_repeat.
 CC Pfam: PF01011; Bacterial_POO_7.
 CC PROSITE: PS00363; BACTERIAL_POO_1; 1.
 CC PROSITE: PS00364; BACTERIAL_POO_2; 1.
 CC Oxidoreductase; POO; Signal; Methanol utilization; Periplasmic.
 CC SIGNAL 1 28
 CC CHAIN 29 626 METHANOL DEHYDROGENASE SUBUNIT 1.
 CC DISULFID 130 131 BY SIMILARITY.
 CC DISULFID 413 442 BY SIMILARITY.
 CC ACT_SITE 330 330 BASE (POTENTIAL).
 CC SEQUENCE 626 AA; 68677 MW; 8768f6b8371e5dfe CRC64;

Query Match 15.7%; Score 430.5; DB 1; Length 626;
 Best Local Similarity 24.28; Pred. No. 1.4e-24;
 Matches 148; Conservative 91; Mismatches 208; Indels 165; Gaps 23;

DB 10 SAGALALIA-AP-AFAQVPTVDELAMPAGE-WISYQNOENRHSPLQTITENVGO 66
 7 SVSALAMALAPALISSAAYANDKVELSKSDNMVWVGKNSNNSELKQVKNVVKO 66
 67 LQLVN-ARGNPGKQVQVPLIHGVMY-----LANPGDV----- 99
 67 LRPATFSTGLNHH-EGAPLYVDGKATVHSFNNTFALDLDPRGILLWQDKPQNPAA 125
 100 -----IQALDAKTGDLWEHRRQLPNATLN 125
 DB 126 RAVACODLVNRLAYWPGDKTPALILKTQLDRHVVALMAGETGVWK----- 173
 126 SGEPTPGKAIYANGVYAGSTCOYSR-----GC-----FVSGHDSATGEEL 168
 174 -----VENSIDIVGSTLTIAPIYVVKQYIISGSGALEGRVGLTAYDVKTGGOV 222
 169 WRNYFI-----PRAGEE--GDETWNNDYEARMTGA---WGQITYPVT 207
 223 WRATAGPPKDLADDDFVKNNAHGGKGLGATWEGD---AMKIGGGTNGWATYADPGT 279
 DB 208 NLVHSGTAVGPASSETQRTGCTGXTNTRAVAPRDGELVWRHQLPRNMWQECTFE 267
 280 NLIVFGTGNPAPNWTMR--PGDNKW-TWTFIRGADPDGEAKFGYQKTPHEMDYAGVNV 336
 268 MMTVNDVQVPTSTEMEGIGSINPNAATGERVYLVGPCKTGTWMPDAETGEFLMAR---D 324
 337 MM-----PSECKD-----KDGKTRKLLTPHNRNGIYYTLDRDGLALVSNKLLDD 380
 DB 325 TNYQNMIESIDENGIYTVNEDALIKELD-VEVDNCPITFLGSDMWSALNDSGIFPL 383
 381 T--VNFVKTLDTLKGQVPRDPEYGTMRDLADVCPASAGYHNOCHDSIDPRRELFEFGI 438
 DB 384 NNVCYDMANVDOEFTSMDVYNTSNVTKLPCK-----DMIGRIDALIDISTGRLMSVER 437
 DB 439 NHIDMDPEPMLYRAGQFVGAITAMMYPGPKDRQNTYGLQIKAYNAVITSGYWKKE 498
 DB 438 AAANYSPVLTGGGVLENGSTDRYPALSOETGETIMQRLATVAGSQAISLEVNGMY 497
 DB 499 RFVAVGGLTATAGDLYVYGTLDYLAKARSDTDLMKRKIPSGAIGVPMYTHKGTOIV 558
 DB 498 ALAGGVSYSGSG 509

DB 559 AL-----YYGVG 565

RESULT 9
 EXAA_PSEAE STANDARD; PRT; 623 AA.
 ID EXAA_PSEAE
 AC 092447;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Quinoprotein ethanol dehydrogenase precursor (EC 1.1.99.-) (OEHD).
 GN EXAA OR PA1982.
 OS Pseudomonas aeruginosa.
 OC Bacteria: Proteobacteria; gamma subdivision; Pseudomonadaceae;
 OC Pseudomonas
 OC NCBI_TaxID=287;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ATCC 17933;
 RX MEDLINE=99041560; PubMed=9826187;
 RA Diehl A., Wintzingerode F., Goerlich H.;
 RT "Quinoprotein ethanol dehydrogenase of Pseudomonas aeruginosa is a
 RT homodimer: sequence of the gene and deduced structural properties of
 RT the enzyme";
 RL Eur. J. Biochem. 257:409-419(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ATCC 15692 / PA01;
 RX MEDLINE=20437337; PubMed=10984043;
 RA Stover C.K., Pham X.-O.T., Erwin A.L., Mizoguchi S.D., Warrenner P.,
 RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
 RA Gierdy R.L., Goltz L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
 RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
 RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
 RA Reizer J., Salier M.H., Hancock R.E.W., Lory S., Olson M.V.;
 RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
 RT opportunistic pathogen";
 RL Nature 406:959-964(2000).
 RN [3]
 RP SEQUENCE OF 1-52 FROM N.A.
 RC STRAIN-ATCC 17933;
 RX MEDLINE=99173751; PubMed=10075429;
 RA Schobert M., Goerlich H.;
 RT "Cytochrome c550 is an essential component of the quinoprotein ethanol
 RT oxidation system in Pseudomonas aeruginosa: cloning and sequencing of
 RT the genes encoding cytochrome c550 and an adjacent acetaldehyde
 RT dehydrogenase";
 RL Microbiology 145:471-481(1999).
 RN [4]
 RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF 35-616.
 RC STRAIN-ATCC 17933;
 RX MEDLINE=20202376; PubMed=10736230;
 RA Kettel T., Diehl A., Knaute T., Stezowski J.J., Hoehne W.,
 RA Goerlich H.;
 RT "X-ray structure of the quinoprotein ethanol dehydrogenase from
 RT Pseudomonas aeruginosa: basis of substrate specificity";
 RL J. Mol. Biol. 297:961-974(2000).
 CC -1- FUNCTION: OXIDISE PRIMARY ALCOHOLS AND ALSO ACTS ON SECONDARY
 CC ALCOHOL, BUT NOT HIGHLY ACTIVE ON METHANOL.
 CC -1- COFACTOR: POO AND CALCIUM.
 CC -1- SUBUNIT: HOMODIMER.
 CC -1- SUBCELLULAR LOCATION: PERIPLASMIC.
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POO DEHYDROGENASE FAMILY.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL: AJ009858; CA00896.1; -
 DR EMBL: AE004624; AAC05370.1; -
 DR EMBL: AF068264; AAC79657.1; -
 DR PDB: 1F1G; 30-AUG-00.
 DR InterPro: IPR001479; Bac_PQQ.
 DR InterPro: IPR002372; Bac_PQQ_repeat.
 DR Pfam: PF01011; Bacterial_PQQ_7.
 DR PROSITE: PS00363; BACTERIAL_PQQ_1; FALSE_NEG.
 DR PROSITE: PS00364; BACTERIAL_PQQ_2; FALSE_NEG.
 DR Oxidoreductase; PQQ; Periplasmic; Signal; Calcium; 3D-structure;
 KM Complete proteome.
 FT SIGNAL 1 34
 FT CHAIN 35 623 QUTNOPROTEIN ETHANOL DEHYDROGENASE.
 FT DISULFID 139 140
 SO SEQUENCE 623 AA; 68123 MW; 32DDE5DF20B291D6 CRC64;

Query Match 15.2%; Score 417; DB 1; Length 623;
 Best Local Similarity 23.1%; Pred. No. 1.4e-23;
 Matches 140; Conservative 95; Mismatches 247; Indels 124; Gaps 20;

QY 3 PSLIMASAGALAL--LAAPAFQVTPYDELLAN--PPAGEWISYGCNGENRHSPLT 57
 DB 9 PACILRPSLHCLAFVALGSGAALAKDVTWEDIANDKTGDVLYQNGTHAQRWSPK 68
 QY 58 QITTEWGOQLVWARGM---QPGVQVTPLIHSGVMTLANPGDVIAIDAKTGDLIME 113
 DB 69 QVANANVFKLTPMSSYSFGDEKORQ--ESQAIYSDGVIVYVYASRFLPAIDAKTKRLMT 127
 QY 114 HRRLP-----NIATLNS----- 126
 DB 128 YNRLRDLRPPCODVYVNRRAIYGVKVFPGTIDASVVALNKTKGVVKKRFAHDGAGYT 187
 QY 127 -PEEPRGAIYANG---VIVAGSTCOYSPFGCF--VSGHDSATGEIMRNYEIPR-- 176
 DB 188 MGCAPF---IYKDGKTGKVLIIHGSS--GDEFVVGRLFARDPDTGSEIMRPFVEGCM 241
 QY 177 -----AGEGDETMGNDEYA-----RMTGA---WGQITYPVTNLVHYGSTAV 217
 DB 242 GRUNGKDSVTVDYKAPSPDPDRNSPTGKVESWSHGGAPQASAPFAETWTIIVGAGNP 301
 QY 218 GPASRTQRTGG-----TLGTNTRFAPVPPDGEIYWRHQTLPRIINWDECFEEMAVT 271
 DB 302 GPWNWARTAKGNPHDYSLY--TSGQYGVDPSSGEVYKMFQHPNDAMDFSGNNELVLF 360
 QY 272 NVDVQPTSTEMEGLQSIINN-----AATGERVLTGVCKTGTW--QFDAETGEFLWARD 324
 DB 361 DYKAKDGKIVATTAHADRNKGFYVDRSNGKIQNAFPFVNITWASHIDLKTG----RP 415
 QY 325 TNYQNMIESIDENGIVTVNEDAILKELDEYDVCPFLGGGRDPSALANDSGIYFPLN 384
 DB 416 VEREGORPPLPEPG-----QKHGKAVEVSPPLGSKNMNPMAYVSDDTGLFYVPAN 465
 QY 385 NVCYDMAVDOEFTSMQVYNTSNVTKLPKGDIMIGRIDAIIDISTGRTLMSVERAANYS 444
 DB 466 HMKEDYMTVEEVSYTKGSAYLGMFRIRKMYDDHVGSLRANDVSGKVMYHMKHEHLYMG 525
 QY 445 VLTSGGVLENGSDRYFRALISOETGETLMTQRLATYASQAIYSEVNGQVYALAGGV 504
 DB 526 VLAIRAGNLVFTGTGDTFKAFDAKSGKELMKFQTGSIYSPITWEDQEOYLGVY--V 582
 QY 505 SYGSGL 510
 DB 583 GYGAV 588

RESULT 10
 DHML_METME STANDARD; PRT; 571 AA.
 AC P38539;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)

DE Methanol dehydrogenase subunit 1 (EC 1.1.99.8) (MDH large alpha
 DE subunit) (MEDH).
 OS Methylophilus methylotrophus (Bacterium W3A1).
 OC Bacteria; Proteobacteria; beta subdivision; Methylophilus group;
 CC Methylophilus.
 OX NCBI_Taxid=17;
 RN [1]
 RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS).
 RX MEDLINE=94059969; PubMed=8241148;
 RA White S., Boyd G., Mathews F.S., Xia Z.-X., Dai W.-W., Zhang Y.-F.,
 RA Davidson V.L.;
 RT "The active site structure of the calcium-containing quinoprotein
 RT methanol dehydrogenase.";
 RL Biochemistry 32:12935-12958(1993).
 RN [2]
 RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS).
 RX MEDLINE=93054513; PubMed=1331050;
 RA Xia Z.-X., Dai W.-W., Xiong J.-P., Hao Z.-P., Davidson V.L., White S.,
 RA Mathews F.S.;
 RT "The three-dimensional structures of methanol dehydrogenase from two
 RT methylophilic bacteria at 2.6-A resolution.";
 RL J. Biol. Chem. 267:22289-22297(1992).
 CC -1- CATALYTIC ACTIVITY: A primary alcohol + acceptor - an aldehyde +
 CC reduced acceptor.
 CC -1- COFACTOR: TWO MOLECULES OF PQQ AND TWO MOLECULES OF CALCIUM
 CC PER TETRAMER.
 CC -1- SUBUNIT: THE HOLOENZYME MDH HAS AN ALPHA-2/BETA-2 CONFIGURATION
 CC OF A LARGE ALPHA AND A SMALL BETA SUBUNIT.
 CC -1- SUBCELLULAR LOCATION: PERIPLASMIC, BUT ASSOCIATED WITH INNER
 CC MEMBRANE.
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL PQQ DEHYDROGENASE FAMILY.
 CC DR PDB: 4AAH; 08-DEC-96.
 DR InterPro: IPR001479; Bac_PQQ.
 DR InterPro: IPR002372; Bac_PQQ_repeat.
 DR Pfam: PF01011; Bacterial_PQQ_7.
 DR PROSITE: PS00363; BACTERIAL_PQQ_1; FALSE_NEG.
 DR PROSITE: PS00364; BACTERIAL_PQQ_2; 1.
 KW Oxidoreductase; PQQ; Methanol utilization; Periplasmic; 3D-structure;
 KW Calcium.
 FT DISULFID 103 104
 FT DISULFID 379 408
 FT ACT_SITE 297 297 BASE (POTENTIAL).
 FT HELIX 2 9
 FT TURN 11 12
 FT STRAND 14 14
 FT TURN 17 18
 FT STRAND 21 22
 FT TURN 26 27
 FT STRAND 34 36
 FT TURN 37 39
 FT STRAND 41 47
 FT STRAND 59 61
 FT TURN 62 63
 FT STRAND 64 68
 FT TURN 71 73
 FT STRAND 75 79
 FT TURN 80 81
 FT TURN 83 84
 FT STRAND 86 90
 FT HELIX 96 101
 FT TURN 103 104
 FT STRAND 112 114
 FT TURN 115 116
 FT STRAND 117 121
 FT TURN 123 124
 FT STRAND 126 131
 FT TURN 132 134
 FT STRAND 137 142
 FT TURN 146 148
 FT HELIX 149 149
 FT TURN 151 151
 FT STRAND 157 159
 FT TURN 160 161

Query Match	14.2%	Score 389.5	DB 1.	Length 571
Best Local Similarity 25.0%		Pred. No. 1.e-21		
Matches 139		Conservative 84	Mismatches 213	Indels 119
Gaps				24
QY	38	AGRWISYSGNOENRHSPLTQITTEWGLQVLW--ARGMOPGVQVTPLIHDGVMTL--	93	
DB	11	AGAMPATGGYYSQHNHSPLAQINKSNVKNRAKMSFSGVINGH-EGAPVLDGMVHVS	69	
QY	94	ANGGDVYQALDAKGTGLIWEHR-RQLPNATLNSGEPTRGMA-----IV---ANGVIYA	144	
DB	70	AFRNNTYALNLNDPGKTIWQHKKPDASTKAVMCDVDRLAYAGAQIYKQKANGHLLA	129	
QY	145	-----GSTCQISPF-----GC-----FVSGHDSATGEEL	168	
DB	130	LDAKGTIMWEVECDPKVGSSTLTAQPFVAKDTLIMGSGSAELGRCGAVNAFEDLKTSELK	189	
QY	169	WR-----NYETPRAGE--EGDETWGNLYEARWMTGA---WGOITYPDVT	207	
DB	190	WRAPATGSDSVLADKDFNSANPHGQFELGKTIWEGD---AMKIGGATMGWYAYDPKL	246	
QY	208	NLVYGSTAVGPASETQGTGG-----TLYSTNTRFAVRPDTEGTEIYWRHQTLPDRWDQ	262	
DB	247	NLFYSGSGNPAPANNETMR--PGDNKKTMTIWRDL-----DTGAKKGYOKTPTHDMDP	298	
QY	263	ECTIEEMAVTVNDQV--STEMEGQS--INNNAATGERHVLTVGPKCTGMPOFDAETGFL	320	
DB	299	AGVQWMLTD--QPVNAKMTPLSHIDKN-----GLTYLNNENGLI	339	
QY	321	WARDTN-YQNMIESIDENGIVATYNEDALIKELDVE-YDVCPTFLGGRDWPSPALNPSGI	378	
DB	340	VAEKVDPAVNVFKVYDLKTGTVPDEFPATRMDHKGTINICPSAMFHNQGVSDVSPESRT	399	
QY	379	YFIPLVNVCYDMAVNDQEFTSMDVYVTSNVTLP----PGKMDGRDAIDISTGRFLMS	434	
DB	400	LYAALNHICMDMEFEMFLPYRAGQFEGATLATMYPGPNCTKKEMQITPAFDITGKAKMT	459	
QY	435	VERNAANVSPVLSGGGVLFGNGSTDRFPALSOETGEITLQRLTAVASQGLSYEDVGM	494	
DB	460	KWEKPAWAGTLLTKYKGVVYATLLDGLYKALNDKCKEILMNNKMSGGIGSPMTYSFRKG	519	
QY	495	QYVAIAGGVSYSGS 509		
DB	520	QYI-----GSMYGVG 529		
RESULT	11			
DHG_ECOLI	STANDARD;	PRT;	796 AA.	
AC	P15877;			
DT	01-APR-1990 (Rel. 14, Created)			
DT	01-NOV-1987 (Rel. 35, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Glucose dehydrogenase [pyrroloquinoline-quinone] (EC 1.1.99.17).			
GN	GCD OR B0124.			
OS	Escherichia coli.			
OC	Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;			
OC	Escherichia.			
OX	NCBI_TaxID=562;			
RN	[1]			
RN	SEQUENCE FROM N.A.			
RC	STRAIN-K12;			
RX	MEDLINE=91035240; PubMed=2228962;			
RA	Cleon-Jansen A.-M., Goosen N., Fayet O., van de Putte P.;			
RT	"Cloning, mapping, and sequencing of the gene encoding Escherichia			
RL	coli quinoprotein glucose dehydrogenase.";			
RL	J. Bacteriol. 172:6308-6315(1990).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-K12 / W3110;			
RX	MEDLINE=993123180; PubMed=8419307;			
RA	Yamada M., Asaoka S., Saiter M.H. Jr., Yamada Y.;			
RT	Characterization of the gcd gene from Escherichia coli K-12 W3110			

RT and regulation of its expression.";
 RL J. Bacteriol. 175:568-571(1993).
 [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-K12 / W3110.
 RX MEDLINE-94261430; PubMed-8202364;
 RA Fujita N., Mori H., Yura T., Ishihama A.;
 RT "Systematic sequencing of the Escherichia coli genome: analysis of
 the 2.4-4.1 min (110,917-193,643 bp) region.";
 RL Nucleic Acids Res. 22:1637-1639(1994).
 [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN-K12 / MG1655;
 RX MEDLINE-97426617; PubMed-9278503;
 RA Blatter F.R., Plunkett G., Ili, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Siao Y.;
 RT "The complete genome sequence of Escherichia coli K-12.";
 RL Science 277:1453-1474(1997).
 [5]
 RN TOPOLOGY.
 RP MEDLINE-93286127; PubMed-8509415;
 RX Yamada M., Sumi K., Matsushita K., Adachi O., Yamada Y.;
 RT "Topological analysis of quinoprotein glucose dehydrogenase of Escherichia
 coli modelled on that of methanol dehydrogenase from Methylobacterium
 extorquens.";
 RL Biochem. J. 312:679-685(1995).
 CC -1- FUNCTION: GDH IS PROBABLY INVOLVED IN ENERGY CONSERVATION RATHER
 CC THAN IN SUGAR METABOLISM.
 CC -1- CATALYTIC ACTIVITY: D-glucose + acceptor = D-glucono-1,5-lactone +
 CC reduced acceptor.
 CC -1- COFACTOR: POO.
 CC -1- SUBUNIT: MONOMER.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane.
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POO DEHYDROGENASE FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

 DR EMBL; X51333; CAA35706.1; -
 DR EMBL; D12651; BAA02174.1; -
 DR EMBL; D26562; CAB20298.1; -
 DR EMBL; AE000122; AAC73235.1; -
 DR PIR; J01017; J01017.
 DR HSSP; P38339; AAAH.
 DR EcoGene; EGI0369; gcd.
 DR InterPro; IPR001479; Bac_POO.
 DR InterPro; IPR002372; Bac_POO_repeat.
 DR Pfam; PF01011; Bacterial_POO_7.
 DR PROSITE; PS00363; BACTERIAL_POO_1; 1.
 DR PROSITE; PS00364; BACTERIAL_POO_2; 1.
 KW Oxidoreductase; POO; Transmembrane; Inner membrane; Periplasmic;
 KW Complete proteome.
 FT DOMAIN 1 10 CYTOPLASMIC (PROBABLE).
 FT TRANSMEM 11 37 PROBABLE.
 FT DOMAIN 38 40 PERIPLASMIC (PROBABLE).
 FT TRANSMEM 41 58 PROBABLE.
 FT DOMAIN 59 62 CYTOPLASMIC (PROBABLE).
 FT TRANSMEM 63 81 PROBABLE.
 FT DOMAIN 82 95 PERIPLASMIC (PROBABLE).

FT TRANSMEM 96 110 PROBABLE.
 FT DOMAIN 111 118 CYTOPLASMIC (PROBABLE).
 FT TRANSMEM 119 141 PROBABLE.
 FT DOMAIN 142 141 PERIPLASMIC (PROBABLE).
 FT ACT_SITE 466 466 BASE (PROBABLE).
 FT CONFLICT 59 59 R -> L (IN REF. 1).
 FT CONFLICT 149 156 TUSADATP -> HIKRCHT (IN REF. 1 AND 2).
 FT CONFLICT 193 193 N -> K (IN REF. 1 AND 2).
 FT CONFLICT 666 666 Q -> H (IN REF. 2).
 SQ SEQUENCE 796 AA; 86747 MW; D9EDC705A12894E9 CRC64;
 Query Match 12.7%; Score 349.5; DB 1; Length 796;
 Best Local Similarity 22.8%; Pred. No. 2.1e-18;
 Matches 161; Conservative 73; Mismatches 209; Indels 263; Gaps 30;
 QY 6 LLMAS-----ACALALAPAFVAVTPTVDILANPPAGENISYGCNQENRRSPLTQ 58
 DB 135 LTMAGFNDPQEIINGTILSADATPAEA-ISPVAQ-----DMPACGRNQGQGRSPKQ 185
 QY 59 ITTENVGQQLVMA-----RQKQPKV--QVPLIHGKVIYLANPGDVIQADAKTG 109
 DB 186 INADVNHLKAWERTGDYKQPNDEGELTNEVTPIKVGDITLYCTAHQRLFDALDAAGK 245
 QY 110 LIMEHRQLPNIATLNSFG-----
 DB 246 EKHVTD---PELKTNESFQVHTCRGVSYHAKAETASPEVMADCPRIITLPVNDRLAI 302
 QY 129 -----EPTRGMAIVANGVIYVAGS-TCQYS--PFG 154
 DB 303 NAENKLCETEFANKGVNLQSNMPTKPGLYEPTSPITTDKTIYAGSVYDNFSTRETS 362
 QY 155 CFSVGHSAAGELMKNYLPRA-----GEGDETCNDYEARMWGANGQITYPEVT 207
 DB 363 GVIRGFVNTGELMA--PDGAKDPNALPSDEHTFTNS-----PNSAPAYADAKL 413
 QY 208 NLVHGSTAVGPASEFQRTGPGTIGTNTFPAVPDPDGEIYVWRHQTLPNNMQDECFEE 267
 DB 414 DIVYLPNGVTPPIWNGGNTPEDEKASSI-LALNATGKLAMSTQYVHHLMDMD---- 468
 QY 268 MMYTNDVOPSTEMESLOSINPNAATGERRVLTGVCPTGTIMQDPAETGFL----- 320
 DB 469 ----LPAQPT-----LAQITVN--GQKVPVIVAPAKGNIFVLDNRNGELVYAPAEKP 515
 QY 321 ----WADTNYQNMIIESIDENGI-----VTYNEDAILKEL--DVEYD----- 356
 DB 516 VPGAAKGDIVYTPQPSLSRPTKDSGADMGATMFDLYCVMFHOMRYGIGFTPP 575
 QY 357 ----VCEFTFLGGRDMPSAALNP-----
 DB 576 SEQGLVFPNGNLGMFEMGWSISVDPAREVALINPMALPVSUKLIRPGONPMEQKDAKT 635
 QY 375 --DSGI-----YFLPLNNVCYDMMAVDQFTSMQVYNTSVTKLDPGKDMIGRIDAI 425
 DB 636 GREGSIQPGYGVYGVTLN-----PFLS-----PGLPCKQKQAMGYISALD 676
 QY 426 ISTGRTLVSEVERAANYS-----PV-----LSTGGGVYFNGST--DRYRRL 465
 DB 677 LKTNEYVMKKRIGTPODSMPFMPVPVPPNMGMPMLGSPISTAGNVLPFIAATADNYLRAY 736
 QY 466 SOETGETLMQRLATVASQA--ISYEVDGQGYAI-AGGVSYS 508
 DB 737 NMSGKELMQLGRLP--AGGQATPMTEYVNGKQYVVISAGHGSGTCT 780
 RESULT 12
 DHGA_ACICA STANDARD; PRT; 801 AA.
 AC P05465;
 DT 01-NOV-1988 (Rel. 09, Created)
 DT 01-NOV-1988 (Rel. 09, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Glucose dehydrogenase-A [pyrroloquinoline-quinone] precursor


```

DE (EC 1.1.99.17) (Quinoprotein glucose DH) (GDH-A).
GN GDH.
OS Acinetobacter calcoaceticus.
OC Bacteria; Proteobacteria; gamma subdivision; Moraxellaceae;
OC Acetobacter.
OX NCBI_TaxID=471.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-LMD 79.41;
RA MEDLINE=88289368; PubMed=3399393;
RT Cleon-Jansen A.-M., Goosen N., Odle G., van de Putte P.;
RT "Nucleotide sequence of the gene coding for quinoprotein glucose
RT dehydrogenase from Acinetobacter calcoaceticus.";
RL Nucleic Acids Res. 16:6228-6228(1988).
CC -1- CATALYTIC ACTIVITY: D-glucose + acceptor = D-glucono-1,5-lactone +
CC reduced acceptor.
CC -1- COFACTOR: POO.
CC -1- SUBUNIT: MONOMER.
CC -1- SUBCELLULAR LOCATION: Periplasmic.
CC -1- MISCELLANEOUS: ACINETOBACTER CALCOACETICUS CONTAINS TWO DIFFERENT
CC POO DEPENDENT ENZYMES WITH GDH ACTIVITY. GDH-A PREFERENCES
CC 2-DEOXYGLUCOSE AS SUBSTRATE, THE SPECIFIC SUBSTRATES FOR GDH-B ARE
CC DISACCHARIDES.
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POO DEHYDROGENASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X07235; CAA30222.1; -
DR PIR: S00943; S00943.
DR InterPro: IPR001479; Bac_POO.
DR InterPro: IPR002372; Bac_POO_repeat.
DR Pfam: PF01011; Bacterial_POO_7.
DR PROSITE: PS00363; BACTERIAL_POO_1; 1.
DR PROSITE: PS00364; BACTERIAL_POO_2; 1.
DR OXidoreductase; POO; Transmembrane; Periplasmic; Signal.
KW SIGNAL
FT 1 33
FT CHAIN 34 801
FT FT
FT TRANSMEM 39 55
FT TRANSMEM 59 79
FT TRANSMEM 94 108
FT TRANSMEM 119 138
FT ACT_SITE 471 471
FT SEQUENCE 801 AA; 86956 MW; 2F4EA222EB5450D8 CRC64;

Query Match 12.0%; Score 329; DB 1; Length 801;
Best Local Similarity 21.9%; Pred. No. 7, 2e-17; Indels 264; Gaps 33;
Matches 152; Conservative 81; Mismatches 197;

OY 20 PAFQVPTVTEDELANPAGEMISYQGNQENYRHSPLTQTTEENVGOLQVWARGQPK 79
DB 152 PETAQAVGVAF-----SMPAYGRTQAGVRSPLKQINDQNVADLKAVMT--LRTGD 202
OY 80 V-----QVPLIHGVMTLANPGDVIAIDAKTDLIW-----EH-- 114
DB 203 IKTNDGSETNQVTPIKGNNMFCITAHOOIAIDPATKEKREDDPKLTKDSFQHLT 262
OY 115 -----RROLP-----NIATLNS-----FGE-- 129
DB 263 CRGWYTYANTTETPATSLSKSSSQCRKRYVPVNDGRIVAVNADGKACTDGGNG 322
OY 130 -----PTRGMATVANGVIVAGS-TCQYS-----PFGCFVSGHDSATGEEL 168
DB 322 QVNIQEFMPVAVPGGYNPTSGIVTGSVVVLAGSVTDNYSNKEPSG-VIRGYVNTGKLL 381
OY 165 WNYITPRAGE-----EGDETGNQDYEARMTGAMQOIYIDPVTNLVHGSAVAVPASE 222

```

```

DB 382 W--VFDTGAADBNAMPGSETEFEHNS-----PNAMAPLADAKLDIV-VYPTGV----- 427
OY 223 TORGPFGSLY-----TNTFAVRPDGTGEIVRROTLDKRDMDQDCEFEAMVTN 272
DB 428 ---GTP--DIWGDRTETLKERYANSMLAINASTGKLVNFOTTHHLDKMDVPSQSLAD 482
OY 273 VDVQSTEMEGLQSLNP--NAATGERRVLTGV-----CKTG--TMMQPAET 316
DB 483 IKKAGQYTPALIVLTGKNAPVLDR--NGQPIVPTREKVPQTVKRGPOTKGEFFSKT 540
OY 317 GEF-----LMA-----RDTYQMMIESIDENGIVTVNEDAIL 348
DB 541 QPFSDLNLPADKLDKDKMDWGMATMDLMCRVSRFLRYVDQILYPPPSNGTL----- 592
OY 349 KELDVEDYDCPFILGGRDPSAALNDS-----GIYFI----- 381
DB 593 -----VPPGNLVGFEMGMSVNPDRQVAVMNPVIGLFPVSRLIPADPNRAQAKAGT 644
OY 382 -----PLNNVYCDMAVVDQFTSMQVYNTSNVTKLPKGMIGRIDAIDISTGRTW--- 433
DB 645 EGVQPMYGVVY-----GVEISAFSLPLGLCPKQAPAGVAVGVDLTHEVWKKR 694
OY 434 -----SYERAANYSPVLSTGGCVLENGT--DRYFRAUSQETGIMOTR 477
DB 695 IGTIRSLPMLFQLPAVKKIGVPGLGSGISTAGNVMFVATQDNTLRANVTNGKKLMEAR 754
OY 478 LATVASGA--ISEYVDGMQYVAI-AGGVSYS 508
DB 755 LP--AGGAATPMYTELNGKQYVIMAGHGSFQT 786

RESULT 13
DHG_GLUOX
ID DHG_GLUOX STANDARD; PRT; 808 AA.
AC P27175;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Glucose dehydrogenase [pyrroloquinoline-quinone] precursor
DE (EC 1.1.99.17).
GN GDH.
OS Gluconobacter oxydans (Gluconobacter suboxydans).
OC Bacteria; Proteobacteria; alpha subdivision; Acetobacteraceae;
OC Gluconobacter.
OX NCBI_TaxID=442.
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=92017653; PubMed=1833618;
RA Cleon-Jansen A.-M., Dekker S., van de Putte P., Goosen N.;
RA "A single amino acid substitution changes the substrate specificity
RT of quinoprotein glucose dehydrogenase in Gluconobacter oxydans.";
RL Mol. Gen. Genet. 229:206-212(1991).
RN [2]
RP REVISION TO 213.
RA Goosen N.;
RL Submitted (FEB-1992) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: D-glucose + acceptor = D-glucono-1,5-lactone +
CC reduced acceptor.
CC -1- COFACTOR: POO.
CC -1- SUBCELLULAR LOCATION: OUTSIDE OF THE PERIPLASMIC MEMBRANE.
CC -1- MISCELLANEOUS: THE P1 FORM CAN OXIDIZE ONLY D-GLUCOSE, WHILE THE
CC P2 FORM CAN ALSO OXIDIZE MALTOSE. THE SEQUENCE OF P1 FORM IS SHOWN
CC HERE.
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POO DEHYDROGENASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

```

CC EMBL: X62710; CAA44594.1; ALT_SEQ.
DR PIR: S17716; OPKEX.
DR InterPro: IPR001479; Bac_PQQ.
DR InterPro: IPR002372; Bac_PQQ_repeat.
DR Pfam: PF01011; Bacterial_PQQ_7.
DR PROSITE: PS00363; BACTERIAL_PQQ_1; 1.
DR PROSITE: PS00364; BACTERIAL_PQQ_2; 1.
KW Oxidoreductase; PQQ; Transmembrane; Periplasmic; Signal.
FT SIGNAL 1 33
FT CHAIN 34 808
FT TRANSMEM 35 54
FT TRANSMEM 59 76
FT TRANSMEM 94 108
FT TRANSMEM 123 138
FT ACT SITE 470 470
FT VARIANT 788 788
FT SEQUENCE 808 AA; 87567 MW; 0F4160DA78652445 CRC64;

```

Query Match 10.7% Score 294.5; DB 1; Length 808;
 Best Local Similarity 20.7%; Pred. No. 2.7e-14;
 Matches 145; Conservative 82; Mismatches 234; Indels 241; Gaps 27;

```

QY 13 ALALLAPAPAPQVTPVDELIA-----NPPAGEWISYQONQENYRHSPLQITTT 61
DB 128 AVTLALFASLFTDPHDSIGELPTQIANASPADPNVPASEWHAYGRQADRMSPLNQIA 187
QY 62 ENVGOLDLVW-----ARQMGKV--QVTEPHDSVMTLANPGVYQAIQAKTGDLIW 112
DB 188 TIVSNIAKVAHHIRTKDMANSNDSEQINTEATPLEFNNTLYMCSLHOKLPAVGATGNVWK 247
QY 113 EHRROL---PIATLN---SFGS-PTGM-----AIYANGVIV----- 143
DB 248 VYDEKLDINPFGHLTCRGVSHFETPANAMDSGNAPITDCAKDSLIPVNDRLVEVDAD 307
QY 144 AGSRC-----QSPRECFV----- 157
DB 308 TGKTCSEFGNGNDELRVNPQPTTGGVEPTSPYITDKLIANSATDNGSVKQASGA 367
QY 158 -SGHDSATGEELW---RNYFIPRAGEGDETMGNDYEARMWGMGAMGQTTYPATNIVHY 212
DB 368 TQADVYTGKRWVYFASNDPQRLDESHPRHPSMWIVS-----SIDANILVYI 422
QY 213 GSRAVGPASTQGTGTP---GGLTYGNTNR---AVRPDTGEIYWRHQTLLPRNDQDEC 264
DB 423 PMGV-----GTPDQMGDRTKDSERFAGIYALNADTKLAFYQVYHHDLMDEL 473
QY 265 TPEMAYTNVDPSTMEGLOSINPAATGERVLTGVCKTGTMQDAEG-ETLMAR 323
DB 474 PGPSPSLVDYTGKGTIVPAIYA-----PTKGTGDIYVLDRTGKEIVAP 517
QY 324 DT-----NYONMIESIDENGIYTVNEDAILKELDY-----EYD 356
DB 518 EFTVPOGAAPBGDHTSPQPMQO---LTLRPKNPLNDSDIWGTIFDQFCSIFETLRYE 574
QY 357 -----VCPFFLGDMPSAALNPDGSIYFIFLANNVCDMAVDO----- 395
DB 575 GPTTPSLGSLIFPDGLDFEMWGLAVDPOROVAFANPISLPEVSQLYVPGPNLWPE 634
QY 396 -----EFTSMVYNTSNVTKLPPGKDM-----GRIDMIDISTGRTL 432
DB 635 ENAKGTGSETGLQHNNGIYAVANLHPFLDPVLLPFGIKMPCRTPMWGVASIDITKNKV 694
QY 433 WSTERRAANYT-----PV-----LSTGGGVLF-NGGTRFYRALSETG 470
DB 695 WQHRNGTLLDSMYGSLPIPLPIKIGVPSLGLGSTAGNIGFLTASMDYITRAYNLTG 754
QY 471 ETLMTRLATVASGAISYEVDMQYVAIAGGVSYSGSLMS 512
DB 755 KYLMQDKRLPAGQAPMTYTAINGKOYI-----VYVAGGHNS 790

```

```

RESULT 14
QY1A_ACICA STANDARD; PRT; 809 AA.
AC 05086;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Quinate/shikimate dehydrogenase [pyrroloquinoline-quinone]
DE (EC 1.1.99.25) (NAD(P)-independent quinate dehydrogenase).
GN QY1A.
OS Acinetobacter calcoaceticus.
OC Bacteria; Proteobacteria; gamma subdivision; Moraxellaceae;
OC Acinetobacter.
OX NCBI_TaxID=471;
RN [1]
RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
RC STRAIN=BD413 / ADP1;
RX MEDLINE=95095936; PubMed=8002591;
RA Elsmore D.A., Ornstom L.N.;
RT "The pca-pob supraoperonic cluster of Acinetobacter calcoaceticus
RT contains quia, the structural gene for quinate-shikimate
RT dehydrogenase."
RL J. Bacteriol. 176:7659-7666(1994).
RN [2]
RP SEQUENCE OF 1-20 FROM N.A.
RC STRAIN=BD413 / ADP1;
RX MEDLINE=96011389; PubMed=7592351;
RA Elsmore D.A., Ornstom L.N.;
RT "Unusual ancestry of dehydrogenases associated with quinate catabolism
RT in Acinetobacter calcoaceticus."
RL J. Bacteriol. 177:5971-5978(1995).
CC -1- FUNCTION: CAN ACT EITHER ON QUINATE OR ON SHIKIMATE.
CC -1- CATALYTIC ACTIVITY: Quinate + pyrroloquinoline-quinone = 5-
CC dehydroquininate + reduced pyrroloquinoline-quinone.
CC -1- CATALYTIC ACTIVITY: SHIKIMATE + PYRROLOQUINOLINE-QUINONE = 3-
CC DEHYDROSHIKIMATE + REDUCED PYRROLOQUINOLINE-QUINONE.
CC -1- COFACTOR: PQQ.
CC -1- PATHWAY: QUINIC ACID CATABOLIC PATHWAY: FIRST STEP. THIS PATHWAY
CC ALLOWS GROWTH OF BACTERIA WITH QUINATE BY ITS CONVERSION TO
CC PROCOCAECHEUATE AND SUBSEQUENT METABOLISM BY THE BETA-KETOGLUTARATE
CC PATHWAY.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- INDUCTION: BY PROCOCAECHEUATE.
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL PQQ DEHYDROGENASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: L05770; AAC37161.1; -
DR InterPro: IPR001479; Bac_PQQ.
DR InterPro: IPR002372; Bac_PQQ_repeat.
DR Pfam: PF01011; Bacterial_PQQ_7.
DR PROSITE: PS00363; BACTERIAL_PQQ_1; FALSE NEG.
DR PROSITE: PS00364; BACTERIAL_PQQ_2; FALSE NEG.
KW Oxidoreductase; PQQ; Quinate metabolism; Transmembrane.
FT TRANSMEM 14 34
FT TRANSMEM 41 61
FT TRANSMEM 68 88
FT TRANSMEM 90 110
FT TRANSMEM 127 147
FT SEQUENCE 809 AA; 88196 MW; 71F67CEBEA62BFCB CRC64;

```

Query Match 9.3%; Score 254.5; DB 1; Length 809;
 Best Local Similarity 21.8%; Pred. No. 2.6e-11;
 Matches 148; Conservative 69; Mismatches 249; Indels 213; Gaps 25;


```

OY 422 DAIDISTGRTLSVRAAAN-----YSPV-----LISTGGVLENGGT-DRY 461
DB 667 SAIDIKTRSIAMQVPVGTVOGTGPPGIGKMLPLIPGMPITLGGTSTGGGLVFIAGTODYY 726
OY 462 FRALSOETGETIMOTRIATVAGOAISY--EVDGMOYVAIAGG 503
DB 727 LRAFDSATGKEIMKGRLEPVGSOGGPITYVSHKTGKQYVVISAGG 770

```

Search completed: May 24, 2002, 10:30:56
 Job time: 634 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2002, 10:30:19 ; Search time 93.8 Seconds
(without alignments)
944.279 Million cell updates/sec

Title: WALICK-934-135.PEP

Sequence: 1 MKPYSILMASGALALIAAP.....GMQYVAIAGGVSYSGSLNS 512

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_19:*
1: sp.archaea:*
2: sp.bacteria:*
3: sp.fungi:*
4: sp.human:*
5: sp.invertebrate:*
6: sp.mammal:*
7: sp.mhc:*
8: sp.organelle:*
9: sp.phage:*
10: sp.plant:*
11: sp.rodent:*
12: sp.virus:*
13: sp.vertebrate:*
14: sp.unclassified:*
15: sp.virus:*
16: sp.bacteriophage:*
17: sp.archaeop:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1511	55.1	608	2	Q93RE9 pseudogluc
2	527.5	19.2	698	2	Q9KH03 alcaligenes
3	495	18.1	742	2	053362 acetobacter
4	488	17.8	601	2	Q9EYH8 rhizobium m
5	487	17.8	601	16	Q92MY9 rhizobium m
6	485.5	17.7	708	2	Q46444 comamonas t
7	482.5	17.6	695	2	Q9F902 pseudomonas
8	470.5	17.2	601	2	P71509 methylobact
9	466.5	17.0	629	2	Q9A048 pseudomonas
10	462.5	16.9	691	2	Q9AP95 pseudomonas
11	428.5	15.6	633	2	Q24759 hyphomicrob
12	420	15.5	623	2	Q9AGW3 pseudomonas
13	419.5	15.3	599	2	Q9L935 methylovoru
14	419.5	15.3	695	2	Q934G0 pseudomonas
15	392.5	14.3	573	2	Q95940 methylophil
16	315.5	11.5	790	2	Q9X2S5 pantoea cit

17	290	10.6	785	16	Q98KE6 rhizobium l
18	287.5	10.5	786	2	P95466 pantoea cit
19	279	10.2	803	16	Q91115 pseudomonas
20	263	9.6	777	16	Q92RB3 rhizobium m
21	234.5	8.6	644	2	052551 pseudomonas
22	220	8.0	639	2	P77931 acetobacter
23	219	8.0	470	2	Q93266 hyphomicrob
24	208.5	7.6	182	2	Q92697 hyphomicrob
25	203.5	7.4	179	2	Q92697 hyphomicrob
26	200.5	7.3	180	2	Q92700 hyphomicrob
27	200.5	7.3	180	2	Q92615 hyphomicrob
28	200.5	7.3	181	2	Q92692 hyphomicrob
29	199.5	7.3	181	2	Q92703 hyphomicrob
30	197.5	7.2	181	2	Q92621 hyphomicrob
31	196.5	7.2	182	2	Q92706 hyphomicrob
32	195.5	7.1	182	2	Q92701 hyphomicrob
33	193.5	7.1	180	2	Q92612 hyphomicrob
34	191.5	7.0	182	2	Q92693 hyphomicrob
35	191.5	7.0	184	2	Q93894 hyphomicrob
36	190.5	6.9	172	2	Q93882 methylosinu
37	186.5	6.8	172	2	Q93882 methylosinu
38	186.5	6.8	179	2	Q93707 hyphomicrob
39	185.5	6.8	180	2	Q92694 hyphomicrob
40	185.5	6.8	180	2	Q93704 hyphomicrob
41	185	6.7	171	2	Q93884 methylosinu
42	183.5	6.7	179	2	Q93702 hyphomicrob
43	182.5	6.7	184	2	Q92695 hyphomicrob
44	181.5	6.6	185	2	Q918X8 uncultured
45	176.5	6.4	172	2	Q93K58 methanotroph

ALIGNMENTS

RESULT ID	1	PRELIMINARY:	PRT:	608 AA.
Q93RE9	Q93RE9	Q93RE9		
AC	Q93RE9	PRELIMINARY:	PRT:	608 AA.
DT	01-DEC-2001 (TRENBERG, 19, Created)			
DT	01-DEC-2001 (TRENBERG, 19, Last sequence update)			
DT	01-DEC-2001 (TRENBERG, 19, Last annotation update)			
DE	ALCOHOL DEHYDROGENASE.			
GN	ADH.			
OS	Pseudoglucobacter saccharofermentans.			
OC	Bacteriophage Pseudoglucobacter.			
OX	NCBI_TaxID=133921;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-IFO 14464;			
RA	Shibata T., Saito Y.;			
RT	"Alcohol dehydrogenase."			
DR	Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.			
DR	EMBL: AB046580; BAB62258.1;			
SQ	SEQUENCE 608 AA; 65101 MW; 0ACEC97AE11BA570 CRC64;			

Query Match	55.1%	Score 1511;	DB 2;	Length 608;
Best Local Similarity	50.7%	Pred. No. 1.3e-101;		
Matches 295;	Conservative 64;	Mismatches 143;	Indels 80;	Gaps 6;
QY	10	SAGALALIAAPFAFO-----VTPVDELLANPPAGGWTGQNGEN 50		
DB	16	STALIASLSGFAFOHDAANAAEFSKAGQSAIENFCPTADDLGKNPAMPILRGNTYG 75		
QY	51	YRHSPLTITTEBNNGOLQVYARGMGQKGVPTLLHIDGYMTLANPQDVQALDAKTGDL 110		
DB	76	WGYSPLDIDINDKDDGLQVMSRTMRPSNEGAIAVNGVIFLANTNDVQLQALDKGTSL 135		
QY	111	IMEHRLPNIAT-LNSRGEPTGMA----- 135		
DB	136	IMEYRKRLPSASKPTNSIGAAKRSIALFPGDKVYFVSMDFVVALDAKTGLAMETNGQS 195		
QY	136	-----IVANGVIVAGSTCYSPFGCFVSGHDSATGEELMRNYFLPRAGEEDDE 184		

```

Db 196 VEEGVANSSGPIVVDGVYIAGSTQFGSGFYGTCTAEGSEBELMRNTEIPRGEGECDDT 255
QY 165 MGN-DYEARNMTGAMGQITTPYVNLVHGSTAVGPASETRGTPGGTLXGNTREAVRP 243
Db 256 WGGAYERNRMTGAMGQITTPYVNLVHGSTAVGPASETRGTPGGTLXGNTREAVRP 315
QY 244 DTGEIWHNQTLPDRNDMDCECTFEMAVTNVYDOPSTEMEGLOSINPAATGE-RVLTGV 302
Db 316 KTEGVVWHQTLPRDNDMDCECTFEMAVTNVYDOPSTEMEGLOSINPAATGE-RVLTGV 375
QY 303 PCKGTGMQFPAETGEFLMARDNTYONMIESIDENGLVTVNEDALIKELDEYDVCPTFL 362
Db 376 PCKGTGMQFPAETGEFLMARDNTYONMIESIDENGLVTVNEDALIKELDEYDVCPTFL 435
QY 363 GGRMPASALNPDSCGYFIPLNNVCYDMAVDEFTSMQVNTSNTKLPKPGDMIGRID 422
Db 436 GGRMPASALNPDSCGYFIPLNNVCYDMAVDEFTSMQVNTSNTKLPKPGDMIGRID 495
QY 423 AIDISTGRTLMSYERAAANSVPLSTGGVLENGGTDRTFRALSOBTGETLMQTRLATVA 482
Db 496 AIDISTGRTLMSYERAAANSVPLSTGGVLENGGTDRTFRALSOBTGETLMQTRLATVA 555
QY 483 SGQAISEVDGMQVVAIAGSG-----VSGSGLN 511
Db 556 SGYTTSTSIDGRTVAAVSGSLGSGPTFGPTTPVDASGAN 597

```

```

RESULT 2
ID 09KH03 PRELIMINARY: PRT: 698 AA.
AC 09KH03;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, last annotation update):
DE TERAHYDROFURFURYL ALCOHOL DEHYDROGENASE.
OS Alcaligenes eutrophus (Ralstonia eutropha).
OC Bacteria: Proteobacteria: beta subdivision; Ralstonia group.
OC Ralstonia
OC NOBL_TaxID=510;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21122557; PubMed=11222593;
RA Zarrin G., Schrader T., Andreesen J.R.;
RT "Catalytic and Molecular Properties of the Quinolhemoprotein
RT Tetrahydrofurfuryl Alcohol Dehydrogenase from Ralstonia eutropha
RT Strain Bo.,"
RL J. Bacteriol. 183:1954-1960(2001).
DR EMBL: AF273737; AAF86335.1; -.
DR HSSP: Q92437; IFLG.
DR InterPro: IPR002372; Bac_PQO_repeat.
DR InterPro: IPR000345; CytC_heme_bind.
DR InterPro: IPR003088; Cyt_C.
DR Pfam: PF01011; Bacterial_PQO; 6.
DR Pfam: PF00034; cytochrome_c; 1.
DR PROSITE: PS00190; CYTOCHROME_C; UNKNOWN.1.
SQ SEQUENCE 698 AA; 75471 MW; 9C87F180BD1D2FB2 CRC64;

```

```

Query Match 19.28; Score 527.5; DB 2; Length 698;
Best Local Similarity 27.98; Pred. No. 5.4e-30;
Matches 163; Conservative 83; Mismatches 220; Indels 119; Gaps 22;
QY 10 SAGALALLAAPFA--QVTPYTDLANPPRG--EWISYQNDENYHSPLTQITTEVNG 65
Db 14 AAASVALPAPMFGANAAARVDGAIRANEGTNPMSYGYDAETRFSLKLEVDNAGNVR 73
QY 66 QLOIYVARGOPGK-VQVTPLIHDGVVYLANPGDVIOAIADAKGDLIMHRRQLP-NIAR 123
Db 74 NLGLAMSYDLSTRGVEATPLVVDGVVAVKAPMSVVAIDARTGKRLWTYDPOVPRDQAV 133
QY 124 LNSGEPPTRGMAI-----VA 138

```

```

Db 134 KGCDDVNRGVYALYQGVFVGAEDGRLVAIDAAATGKKWEDQITVDRSKSYITGAPRV 193
QY 139 NG-VIAGSTQYSPFCFVSGHDSATGEELMNNYFLP-----RAGEGDEMGWD 188
Db 194 NGVILIGGAGLEVVR-YITADAETGKQKMMYVTPGDPARPFENEMAKAAATW-D 250
QY 189 YEAR-WMTG-----AMGQITVPTNLVHGSTAVGPASETRGTPGG-TLYGNTREAVR 242
Db 251 PSGRVWINGGGGVTWNMTAPDELINLMYITGTNAGCWSRKLSPKGDNLVAASV-VALN 309
QY 243 PDTEIWHNQTLPDRNDMDCECTFEMAVTNVYDOPSTEMEGLOSINPAATGERVLT 300
Db 310 PDTEIWHNQTLPDRNDMDCECTFEMAVTNVYDOPSTEMEGLOSINPAATGERVLT 375
QY 301 GVPCKTGMQFPAETGEFLMARDNTYONMIESIDENGLVTVNEDALIKELDEYDVCPTFL 362
Db 352 HAP-KNGFEVVIDRTNKFISAKNFVDVWASGDKNGRPHVTPQADTSGKRA-----D 406
QY 359 PTFILGGRDMPASALNPDSCGYFIPLNNVCYDMAVDEFTSMQVNTSNTKLPKPGDMIGRID 422
Db 407 PTFILGGRDMPASALNPDSCGYFIPLNNVCYDMAVDEFTSMQVNTSNTKLPKPGDMIGRID 495
QY 408 ---VTKLPKDMIGRIDAIDISTGRTLMSYERAAANSVPLSTGGVLENGGTDRTFRALSOBTGETLMQTRLATVA 482
Db 462 LGMVNAEPFRSKPMGRLLAMPDLAKAVMHRDHGPMNGGLATAGNLYVGTADGRV 521
QY 464 ALSQETGETLMQTRLATVASGQAISEVDGMQVVAIAGSG-----VSGSGLN 511
Db 522 AYHAATGKLMQAPRGSGVVAAPVYTLIDGRQVSAVAGMGVYG 566

```

```

RESULT 3
ID 053362 PRELIMINARY: PRT: 742 AA.
AC 053362; Q44139;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, last sequence update)
DT 01-OCT-2001 (TrEMBLrel. 18, last annotation update)
DE DEHYDROGENASE SUBUNIT OF ALCOHOL DEHYDROGENASE.
OS Acetobacter pasteurianus (Acetobacter turbidans).
OC Bacteria: Proteobacteria: alpha subdivision; Acetobacteraceae;
OC Acetobacter.
OC NOBL_TaxID=438;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94042848; PubMed=8226628;
RA Takemura H., Kondo K., Horinouchi S., Beppu T.;
RT "Induction by ethanol of alcohol dehydrogenase activity in Acetobacter
RT pasteurianus.,"
RL J. Bacteriol. 175:6857-6866(1993).
DR EMBL: D13893; BAA40252.1; -.
DR HSSP: Q92437; IFLG.
DR InterPro: IPR001479; Bac_PQO.
DR InterPro: IPR002372; Bac_PQO_repeat.
DR Pfam: PF01011; Bacterial_PQO; 6.
DR PROSITE: PS00363; BACTERIAL_PQO; 1.
DR PROSITE: PS00364; BACTERIAL_PQO; 2; 1.
SQ SEQUENCE 742 AA; 81680 MW; 031C4E9B6B13EC98 CRC64;

```

```

Query Match 18.18; Score 495; DB 2; Length 742;
Best Local Similarity 27.08; Pred. No. 1.4e-27;
Matches 164; Conservative 82; Mismatches 221; Indels 140; Gaps 22;
QY 11 AGALALLAAPFAQVTPYTDDEL--ANPAGEWISYQNDENYHSPLTQITTEVNG 68
Db 23 AAALPVAIVPARADGQGTGATIIHADHPENMLSYGRYSRYSRPLQDINSNVGDLK 82
QY 69 LVNARGMOPGKVVQ-VTPPLIHGVVYLANPGDVIOAIADAKGDLIMHRRQLP-NIAR 122
Db 83 LAWYITLDTNRGQAEATPLVVDGVVAVKAPMSVVAIDARTGKRLWTYDPOVPRDQAV 133

```

```

QY 123 --TLN-----SRGE-----PTRGMAI 136
DB 143 CDTVNRAGATYNGKQVWGTGFDRLVAADAKTGKRWVENITPADASLGKQSTYDVAVR 202
QY 137 VANGVIVASTGOYSPGCG--FVSGHDSATGEELMRNYFI PRAGEEDG-----182
DB 203 VANGVIVASTGOYSPGCG--FVSGHDSATGEELMRNYFI PRAGEEDG-----182
QY 183 -ETWGNDEYARWMT-----GANGQITTYDEVTNLVHGSTAVGPASETORGTGTLT---233
DB 261 YKTWGP--KGAVYRQGGGSTDVSDLYDVSLLIY---LAVG-----NGSPMYKYRSE 309
QY 234 --GTN-----TRPVRPDTEIYWRHQTLPROWWDECTFEMWVTNVDPQSTEMEGLOSI 287
DB 310 GISSNIFLSIALKPEETGEYVHFOATPMDQDYSQOITLTLDMY-----357
QY 288 NPNAATGE--RRVLTVGPCTGTMOQFADTGEFLWADRTNTONMIESIDE-NGIYTVNED 345
DB 358 -----NGEMRHVIMHAP-KNGFFYVLDATGTEFLAGKRVYQNMANGIDLPLGRPIYND 411
QY 346 AILKELDEYDVCPTFLGGRDMPAALNDSCITFIPLNNVY-----DMMMA 392
DB 412 GLYTLTGKFFYVGPFLGAGNEMGAYSPKTHLYLPAHQIPFGYKNOVGKPPHDAWN 471
QY 393 VDDEFTSMDYNTSNVTKLPKCKMIGRIDAIDISTGRTLSVERAANYSPVLSTGGGV 452
DB 472 VGLDMTKNGLPDTPPE-ARRAYIKDLHGMILLANDPYKMETVWKIDKSPMNGSVLATGDL 530
QY 453 LFNGGDRYFRALSOETGELTMOATLATAVSAQAISYEVDGQYVA-----IAGG 502
DB 531 LFGGLAAGEFHAYDANGSDLYKFDAGSGIILASPMYTVNGQYVAVEVGWGIYPISMG 590
QY 503 GVSYGSG 509
DB 591 GVGRTSG 597

```

RESULT 4

Q92WY8 PRELIMINARY; PRT: 601 AA.

AC Q92WY8; 01-MAR-2001 (Tremblrel. 16, Created)

DT 01-DEC-2001 (Tremblrel. 16, last sequence update)

DT 01-DEC-2001 (Tremblrel. 19, last annotation update)

DE METHANOL DEHYDROGENASE LARGE SUBUNIT-LIKE PROTEIN.

GN KMAP.

OS Rhizobium meliotti (Sinorhizobium meliotti).

OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;

OC Rhizobiaceae; Sinorhizobium.

OX NCBI_TaxID=382;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-RM1021;

RA Fenner B.J., Tiwari R.P., Dilworth M.J.; "Regulation of C₁ assimilation in Sinorhizobium meliotti."

RT Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.

RL EMBL; AF309488; AAC31643.1; "

DR HSSP; P38539; 4AAH.

DR Interpro: IPR002372; Bac_PQQ_repeat.

DR Pfam; PF01011; Bacterial_PQQ_7

DR SEQUENCE 601 AA; 65686 MW; 09F743726B2E2A83 CRC64;

Query Match 17.88; Score 488; DB 2; Length 601;
 Best Local Similarity 26.28; Pred. No. 3.2e-27;
 Matches 157; Conservative 90; Mismatches 215; Indels 138; Gaps 24;

```

QY 14 LALLAAPFAOYTVPTDEL--LANPPAGEMISGQONENRHSPLQITTEENYGOLOLV 71
DB 8 LAIMSIGGAOYAFANDELQKLIIDP--NOMALIGTDVANIARYSKLDQINKDNVGLQVAM 66
QY 72 --ARMGQKQVAPFLINDGVYLANP-GDVYQIAD-AKTGDILWEHR-RQLPNITLNS 126

```

```

DB 67 TESTGLRGH-EGSPLVIGDLMYVHTPPNVTVALDLSDKGQIWKYEPKODNPVIPYC 125
QY 127 FGEPTRGMAI-----VANGVIVAGST-----147
DB 126 CDTVNRAGATYNGKQVWGTGFDRLVAADAKTGKRWVENITPADASLGKQSTYDVAVR 202
QY 148 ---CQYSPGCG--FVSGHDSATGEELMRNYFI PRAGEEDG-----182
DB 186 LVGISGGEFVGAVHTAVSMADGKVLMRGYSKGPDSPLIDPEKTHLGRKFGKDSGLTT 245
QY 185 WGNDEYARWMT-----ANGQITTYDEVTNLVHGSTAVGPASETORGTGTLT---233
DB 246 WEGD-----OMKIGGTTWGYSDPEENLVYTGTPSTWNPQR--PQDNW-SMTIFAR 299
QY 242 RPDTEIYWRHQTLPROWWDECTFEMWVTNVDPQSTEMEGLOSIINPAATGERRVLG 301
DB 300 DVDTGMAKLLXOMTPRDEMDYGVNEMILTGHDGK-----DRKLLTH 343
QY 302 VPCKTGTMOQFADTGEFL-----WADDTYKQNMIE-----SIDENGIV 340
DB 344 FD-RNGFGYTMORVTEGELLVAEKYDPTVMATEVMDPKSDYGRPOVAYQISTEONG--400
QY 341 TVNEDAILKELDEYDVCPTFLGGRDMPAALNDSCITFIPLNNVYCYDMAVDOEFTSM 400
DB 401 ---EDNTT-----GVCRAALGTNDQOPAAISPKTELFYVPTNIVCMDEPFRVSYTAG 451
QY 401 DYNTSNVTKLPKCKM--IGRIDAIDISTGRTLSVERAANYSPVLSTGGGV 452
DB 452 QPYVGAATLSWYPP-KDSHGSGMNFICADNKEGKIKMSIPEPFSVSGALATAGDVFFGT 510
QY 458 TDKYFRALSOETGELTMOATLATAVSAQAISYEVDGQYVA-----IAGG 502
DB 511 LEGYLAADAATSKELRYKRTSPGVIGVNTYARBEQYVAVLSGVGMAIGLAAGLTN 570

```

RESULT 5

Q92WY9 PRELIMINARY; PRT: 601 AA.

AC Q92WY9; 01-DEC-2001 (Tremblrel. 19, Created)

DT 01-DEC-2001 (Tremblrel. 19, last sequence update)

DT 01-DEC-2001 (Tremblrel. 19, last annotation update)

DE PUTATIVE METHANOL DEHYDROGENASE PROTEIN, LARGE SUBUNIT (BC 1.1.99.8).

GN SMH20173.

OS Rhizobium meliotti (Sinorhizobium meliotti).

OG Plasmid pSymb (megaplasmid 2).

OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;

OC Rhizobiaceae; Sinorhizobium.

OX NCBI_TaxID=382;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-1021;

RA MEDLINE-21396508; Pubmed-11481431; Firan T.M., Weidner S., Wong K., Buhrmester J., Chain P., Vorholster F.J., Hernandez-Lucas I., Becker A., Cowie A., Gouzy J., Golding B., Puelher A.; "The complete sequence of the 1.683-kb pSymb megaplasmid from the N2-fixing endosymbiont Sinorhizobium meliotti."

RT Proc. Natl. Acad. Sci. U.S.A. 98:9889-9894(2001).

RL EMBL; AL603642; CAC48573.1; "

KW Oxidoreductase; Plasmid; Hypothetical protein; Complete proteome.

DR SEQUENCE 601 AA; 65759 MW; D73424EFED15ADBFCRC64;

Query Match 17.88; Score 487; DB 16; Length 601;
 Best Local Similarity 26.28; Pred. No. 3.8e-27;
 Matches 157; Conservative 90; Mismatches 215; Indels 138; Gaps 24;

```

QY 14 LALLAAPFAOYTVPTDEL--LANPPAGEMISGQONENRHSPLQITTEENYGOLOLV 71
DB 8 LAIMSIGGAOYAFANDELQKLIIDP--NOMALIGTDVANIARYSKLDQINKDNVGLQVAM 66

```



```

OY 72 -- ARMGPGKQVPTLHIDGVWYLANP-GDVIAID-AKTGDLIMEHR-ROLPIATLNS 126
DB 67 TESTGVLRGH-EGSPVLYGIDLMYHTPEPPNVYALDLSKQOIVMKKEPKODENRVLPMVC 125
OY 127 EGEPRGMAI-----VANGVYAGST-----147
DB 126 CDTNRYGVAVDNKLIFLHQADTIVVADAKTGKIVMSVKNKGDAIKGTNTATVMPYDKI 185
OY 148 ---COYSPFGC--FVSGHDSATGEELWNRNYFI-----PRAGEGDEI 184
DB 166 LVGISGGEFVGRVHTAVASMDGKVLKRGISMGSDSTLLIDPEKTHLKGKPVSGKDSGLTT 245
OY 165 WGNDEANWMTG---AMGQITVPTNLVHSGTAVGPASETGRTGPGTLYGTNTFAY 241
DB 246 WEGD---OMKIGGTTWGMYSYDPEENLYYGTGNTSPMTWPTQR--PGDNRM-SMTIFAR 299
OY 242 RPDGEIYWRHQTLPBPNWMOCEFEEMATNNVQVSTEGEGLSTINPNAATGERVLTG 301
DB 300 DVDIGMAKWLQMPRHDEMDIDGVNEMILTEQOIDGK-----DRKLLTH 343
OY 302 VPCRTGTMQDPAETGEEL-----WARDNYONMIE-----SIDENGIV 340
DB 344 FD-RNGGTYTMDRYTGELVAEKYDPTVNWATEVMDPKSKXGRPOVVAQYSTENG-- 400
OY 341 TVNEDALIKELDYEDYDCPTFLGGRMPKSPALMPDSCITFIPIPLNNGVYMAVDOETFSM 400
DB 401 ---EDTWT-----GVCPALGTRKQDQPAAYSPKTELEVYPNHNKCMDEEPRVSYTAG 451
OY 401 DVYNTSVNTKLPGRKDM---IGRIDADISTGRTLMSVERAANYSPLVSTGGGVLFNG 457
DB 452 QPYGATILSMYPR-KDSHSGMGNFNIAMDNEKIKMSLPEPPSVMGALATAGDVFYGT 510
OY 458 TDRFRALSOPTGETLMQTLATVASGQALSYVDGMQVYAIAG-----GVSYSGSLNS 512
DB 511 LEGILKAVDADATGELRYEFTKPSGVIGNWTVAREGKQVAVVLSGVGMAIGIGLAAGLTN 570

RESULT 6
OY 046444 PRELIMINARY: PRT: 708 AA.
AC 046444;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE QUINOAEMOPROTEIN ETHANOL DEHYDROGENASE TYPE I PRECURSOR
DE (P1.1.99.1) (QH-EDH1).
OS QHEDH.
ON Comamonas testosteroni (Pseudomonas testosteroni).
OC Bacteria; Proteobacteria; beta subdivision; Comamonadaceae; Comamonas.
OX NCBI_TaxID=285;
RN 11
RP SEQUENCE FROM N.A.
RX STRAIN-ATCC 15667;
RX MEDLINE=96184549; PubMed=8654419;
RA Stoetvoegel J., Kraayveld D.E., van Stuils C.A., Jongejan J.A.,
RA de Vries S., Duine J.A.;
RT "Characterization of the gene encoding quinochaemoprotein ethanol
RT dehydrogenase of Comamonas testosteroni";
RL Eur. J. Biochem. 235:690-698(1996).
RN 12
RP SEQUENCE OF 32-54 AND 477-490.
RX STRAIN-ATCC 15667;
RX MEDLINE=95324580; PubMed=7601151;
RA De Jong G.A.H., Geerlof A., Stoetvoegel J., Jongejan J.A., De Vries S.,
RA Duine J.A.;
RT "Quinochaemoprotein ethanol dehydrogenase from Comamonas testosteroni.
RT Purification, characterization, and reconstitution of the apoenzyme
RT with pyroloquinoline quinone analogues.";
RL Eur. J. Biochem. 230:899-905(1996).
RN 13
RP CHARACTERIZATION.
RX MEDLINE=86242113; PubMed=3521592;
RA Groen B.W., van Kleef M.A., Duine J.A.;

```

```

RT "Quinochaemoprotein alcohol dehydrogenase apoenzyme from Pseudomonas
RT testosteroni";
RT Biochem. J. 234:611-615(1986).
RN 14
RP CRYSTALLIZATION.
RX MEDLINE=21536088; PubMed=11679760;
RA Oubrie A., Huizinga E.G., Rozeboom H.J., Kalk K.H., de Jong G.A.H.,
RA Duine J.A., Dijkstra B.W.;
RT "Crystallization of quinochaemoprotein alcohol dehydrogenase from
RT Comamonas testosteroni: crystals with unique optical properties.";
RL Acta Crystallogr. D 57:1732-1734(2001).
CC -1- FUNCTION: CATALYZES THE DYE-LINKED OXIDATION OF PRIMARY ALCOHOLS
CC TO THE CORRESPONDING ALDEHYDES AND THE (SUBSEQUENT) OXIDATION OF
CC THE ALDEHYDES TO CARBOXYLIC ACIDS.
CC -1- COFACTOR: PQQ, HEME, AND CALCIUM.
CC -1- SUBUNIT: MONOMER.
CC -1- SUBCELLULAR LOCATION: PERIPLASMIC (POTENTIAL).
CC -1- INDUCTION: BY ETHANOL AND BUTANOL.
CC -1- SIMILARITY: TO ALCOHOL DEHYDROGENASE (ACCEPTOR) PROTEINS.
DR EMBL: X81880; CA57464.1;
DR HSSP: Q924J7; 1F1G.
DR InterPro: IPR002372; Bac_PQQ_repeat.
DR InterPro: IPR003086; Cyt_C1.
DR InterPro: IPR002329; Cyt_C1C.
DR Pfam: PF01011; Bacterial_PQQ_6.
DR Pfam: PF00034; Cytochrome_c_1.
DR PRINTS: PR00605; CYTOCHROME_C1C.
KW Signal; PQQ; Heme; Calcium; Oxidoreductase; Periplasmic.
FT SIGNAL 1 31
FT CHAIN 32 708
FT TYPE 1.
FT BINDING 635 635 QUINOAEMOPROTEIN ETHANOL DEHYDROGENASE
FT COVALENT 638 638 HEME (COVALENT) (BY SIMILARITY).
FT METAL 639 639 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
SQ SEQUENCE 708 AA; 76822 MW; 99AB54BDD6ACCB3 CRC64;

```

Query Match 17.7%; Score 485.5; DB 2; Length 708;
 Best Local Similarity 25.7%; Pred. No. 6, 2e-27;
 Matches 152; Conservative 86; Mismatches 241; Indels 113; Gaps 20;

```

OY 3 PTSLIWSAGLALLAAPPAPQVTPVT-----DELIANPP-AGEMISYGQDEN 50
DB 12 PGMWMLAACLG--SAAAFQOTGPAQAAAQVHVDDEFIRANARFPDPTGIVYAE 69
OY 51 YRSPLOTITTEVNGOLOVWARGMOPCK-VQVPLHIDGVWYLANPQDVQALDATGCD 109
DB 70 TRISRLDQIMANVAVKDLGLAMSYNLESTRGVEATFVVDGIMTVSASVVAHIDTRGN 129
OY 110 LIWEHRRLP-----NIAT-----LNSFG 128
DB 130 RIWTYPQIDRSTGFGKCCDVNRGVALMKGVYGVGAMDRLLALDADATGKVEWHQNTF- 168
OY 129 EPTRGMAIYANG-----VIYAGSTCOYSPGCGVSGHDSATGEELMRNTFTF----- 175
DB 189 EGKGSGLTTGAPRVPRKGVYIIGKRAEYGVAG-VITAYVDETEERKMRFSVGDPSKP 247
OY 176 ---RAGEGDETFMGNDYEARMWTCG----WGQITVPTNLVHSGTAVGPASETGRT 227
DB 248 FDESKMKRAARFTW--DPGKMWEEAGGGGTMDSMTFDELNTNIVGVGKNSPSHKYRSP 305
OY 228 PGGLTLYGNTRAVPRDTGELVWRHQTLPBPNWMOCEFEEMATNNVQVSTEGEGLST 287
DB 306 KGDMLYLASIVALPDPGKYKHQETPGDNMDYTSFQPIILADIKI----- 353
OY 288 NPNNAATGERRYLVGPCKTGTMQDPAETGEELMARPTNTONNTESIDENG-IYTVNEDA 346
DB 354 ---AGKPKVYLHAP-KNGEFFVLDRTNKFEISAKNFVPMNASGDKHGKPDGI---A 405
OY 347 ILKELDYEDYDCPTFLGGRMPKSPALMPDSCITFIPIPLNNGVYMAVDO-EFT----- 398
DB 406 AARDOSKQDANVPGRYGAHNHNPMSFNPOIGLYLVLPQANVPVNLMDKKWEFNQAGPKRP 465

```

ID	Accession	Score	Length	DB	DB2	Length	DB1	DB2	DB3	DB4	DB5	DB6	DB7	DB8	DB9	DB10	DB11	DB12	DB13	DB14	DB15	DB16	DB17	DB18	DB19	DB20	DB21	DB22	DB23	DB24	DB25	DB26	DB27	DB28	DB29	DB30	DB31	DB32	DB33	DB34	DB35	DB36	DB37	DB38	DB39	DB40	DB41	DB42	DB43	DB44	DB45	DB46	DB47	DB48	DB49	DB50	DB51	DB52	DB53	DB54	DB55	DB56	DB57	DB58	DB59	DB60	DB61	DB62	DB63	DB64	DB65	DB66	DB67	DB68	DB69	DB70	DB71	DB72	DB73	DB74	DB75	DB76	DB77	DB78	DB79	DB80	DB81	DB82	DB83	DB84	DB85	DB86	DB87	DB88	DB89	DB90	DB91	DB92	DB93	DB94	DB95	DB96	DB97	DB98	DB99	DB100	DB101	DB102	DB103	DB104	DB105	DB106	DB107	DB108	DB109	DB110	DB111	DB112	DB113	DB114	DB115	DB116	DB117	DB118	DB119	DB120	DB121	DB122	DB123	DB124	DB125	DB126	DB127	DB128	DB129	DB130	DB131	DB132	DB133	DB134	DB135	DB136	DB137	DB138	DB139	DB140	DB141	DB142	DB143	DB144	DB145	DB146	DB147	DB148	DB149	DB150	DB151	DB152	DB153	DB154	DB155	DB156	DB157	DB158	DB159	DB160	DB161	DB162	DB163	DB164	DB165	DB166	DB167	DB168	DB169	DB170	DB171	DB172	DB173	DB174	DB175	DB176	DB177	DB178	DB179	DB180	DB181	DB182	DB183	DB184	DB185	DB186	DB187	DB188	DB189	DB190	DB191	DB192	DB193	DB194	DB195	DB196	DB197	DB198	DB199	DB200	DB201	DB202	DB203	DB204	DB205	DB206	DB207	DB208	DB209	DB210	DB211	DB212	DB213	DB214	DB215	DB216	DB217	DB218	DB219	DB220	DB221	DB222	DB223	DB224	DB225	DB226	DB227	DB228	DB229	DB230	DB231	DB232	DB233	DB234	DB235	DB236	DB237	DB238	DB239	DB240	DB241	DB242	DB243	DB244	DB245	DB246	DB247	DB248	DB249	DB250	DB251	DB252	DB253	DB254	DB255	DB256	DB257	DB258	DB259	DB260	DB261	DB262	DB263	DB264	DB265	DB266	DB267	DB268	DB269	DB270	DB271	DB272	DB273	DB274	DB275	DB276	DB277	DB278	DB279	DB280	DB281	DB282	DB283	DB284	DB285	DB286	DB287	DB288	DB289	DB290	DB291	DB292	DB293	DB294	DB295	DB296	DB297	DB298	DB299	DB300	DB301	DB302	DB303	DB304	DB305	DB306	DB307	DB308	DB309	DB310	DB311	DB312	DB313	DB314	DB315	DB316	DB317	DB318	DB319	DB320	DB321	DB322	DB323	DB324	DB325	DB326	DB327	DB328	DB329	DB330	DB331	DB332	DB333	DB334	DB335	DB336	DB337	DB338	DB339	DB340	DB341	DB342	DB343	DB344	DB345	DB346	DB347	DB348	DB349	DB350	DB351	DB352	DB353	DB354	DB355	DB356	DB357	DB358	DB359	DB360	DB361	DB362	DB363	DB364	DB365	DB366	DB367	DB368	DB369	DB370	DB371	DB372	DB373	DB374	DB375	DB376	DB377	DB378	DB379	DB380	DB381	DB382	DB383	DB384	DB385	DB386	DB387	DB388	DB389	DB390	DB391	DB392	DB393	DB394	DB395	DB396	DB397	DB398	DB399	DB400	DB401	DB402	DB403	DB404	DB405	DB406	DB407	DB408	DB409	DB410	DB411	DB412	DB413	DB414	DB415	DB416	DB417	DB418	DB419	DB420	DB421	DB422	DB423	DB424	DB425	DB426	DB427	DB428	DB429	DB430	DB431	DB432	DB433	DB434	DB435	DB436	DB437	DB438	DB439	DB440	DB441	DB442	DB443	DB444	DB445	DB446	DB447	DB448	DB449	DB450	DB451	DB452	DB453	DB454	DB455	DB456	DB457	DB458	DB459	DB460	DB461
----	-----------	-------	--------	----	-----	--------	-----	-----	-----	-----	-----	-----	-----	-----	-----	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------

QY 348 LKELDEYD-----VCPTEFGDRMPAALNDPSGIFYIFLPLNNCYDMAVDEFTS 399
 Db 390 VSKSTEDNGDVNSKCPALCTKQOAPASPKTGLTYVPLNHCMDYEPFRVYTP 449
 QY 400 MDVYTSVYTLR-PGR-DMIGRIDAIDISTGRTLSVERAANYSPLYSTGGVLFNGG 457
 Db 450 GQPVYATLSMYPARGSHGGMNFIAWNLQKIKMSNPQSFAMGALATSGDVLFGYT 509
 QY 458 TDRFRALSOETGELTNOTRLATVAASGATSYEDVQOYALNG-----GVSYSGL 510
 Db 510 LEGFLKAVDSKTGKELKFKTPSGITGNVMTYEHKGRHVAALVSGVAGAGL 567

RESULT 9
 QY 09A048 PRELIMINARY; PRT: 629 AA.
 AC 09A048;
 DT 01-JUN-2001 (Tremblrel. 17, Created)
 DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE METHANOL DEHYDROGENASE MXAF.
 GN MXAF.
 OS Methylobacterium nodulans.
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Methylobacterium group; Methylobacterium.
 OX NCBI_TaxID-114616;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ORS2060;
 RX MEDLINE-2056686; PubMed-11114919;
 RA Sy A., Giraud E., Jourand P., Garcia N., Willems A., de Lajudie P.,
 RA Peth Y., Neyra M., Gillis M., Bolvin-Masson C., Dreyfus B.,
 RT Methylobacterium nodulans Bacteria nodulate and fix nitrogen in
 RT symbiosis with legumes.
 RL J. Bacteriol. 183:214-220(2001).
 DR EMBL: AF220764; AAG49450.1;
 DR HSP: P38539; 4A4H.
 DR InterPro: IPR001479; Bac_P00.
 DR InterPro: IPR002372; Bac_P00_repeat.
 DR Pfam: PF01011; Bacterial_P00_1;
 DR Prosite: PS00363; Bacterial_P00_1;
 DR Prosite: PS00364; Bacterial_P00_2;
 DR PROSITE: PS00364; Bacterial_P00_2;
 SO SEQUENCE 629 AA; 68798 MW; 3DB559D8373BB0B2 CRC64;

Query Match 17.0%; Score 466.5; DB 2; Length 629;
 Best Local Similarity 24.9%; Pred. No. 1.3e-25;
 Matches 150; Conservative 98; Mismatches 214; Indels 141; Gaps 22;

QY 4 TSLMASAGALALAAPAFQVPTVTDLL-ANPPAGEWISYGNONENRHSPLTQTTE 62
 Db 11 TGVSYAALALALPLAGPS-----ALANDKLVLSKSDGNWMPGKNYSDNYSKLQINAE 66
 QY 63 NNGQQLVW--ARGMQPKVOYVPLIHGVMY-----LANPGDV----- 99
 Db 67 NVKNLKVMQSFSTGLLNCH-EGAPLVVDGVTYVTSFNNTEFALGDDEPKIIMQKPKQ 125
 QY 100 -----IOAIDAKTGDILWEHRRQIPNT 121
 Db 126 NPAARSVACCLVNRGLAYWPGDCKTSLILKILLDGHVVALNAQGEYTKMLTENDIRY 185
 QY 122 -ATLNSGEPTRGMAIVANGYVAGSTQYSPFCFVSGHDSATGEIMNRYFIPRAGEE 180
 Db 186 GSTL-----TIAPYVVKDKVLISSGAEIGVNG-YLTAVDVTGDKMKRAY-----ATGSD 235
 QY 181 GD-----ETWGNDEYEAAMTGA---WGQIYVDPVNTLVHVGSTA 216
 Db 236 SDLLADENITHNAHYQCKGLGTSTWEGD---AMKIGGCTNMGVYADPGTNTLIYTGCTN 292
 QY 217 VGPASQTRGTPGGLTYGNTREFAVRPTGEIYWRHQTLPROMWDECFEMMYVTVDVQ 276
 Db 293 PAPPNETMR--PGDNKV-TWTFIRARDVDYDEAKFGYQKTPHDEMDYAGVNVML----- 343

QY 277 PSTEMEGLOSINPNAATGERRVLTGPCKTGTWMOFDAETGEFLMAR---DTNYOMIRS 333
 Db 344 -STOKD-----RSKEREKLLTHPDNRNGIYYTIDRTNGDLSHAKIDDY--VAVERT 391
 QY 334 IDENGITYVNDALKELD-VEYDVCPTEFLGDMPSAALNDPSGIFYIFLPLNNCYDMA 392
 Db 392 VDLKSGLVPRDPEYGTGRTMDLARDICPSAMGYNOCHDSYDPERKLFYMGINHCMDPE 451
 QY 393 VDQETSMQVNTSNVTKLPGRK-----DMIGRIDAIDISTGRTLSVERAANYSPLY 446
 Db 452 FMLPYRAGQEFFVAGTLNMYGPRKGDQONNEGLOQIAYATITGKREWEKFAVWGTL 511
 QY 447 STGGVLFNGSTDRYFRALSQETGELTNOTRLATVAASGATSYEDVQOYALNGVSY 506
 Db 512 ATAGNVVFTGLDGLFKARHSDTEGLIMRAKILPSGAIQVPTYTHKGQYVAI-----Y 566
 QY 507 GSG 509
 Db 567 GVG 569

RESULT 10
 QY 09AF95 PRELIMINARY; PRT: 691 AA.
 AC 09AF95;
 DT 01-JUN-2001 (Tremblrel. 17, Created)
 DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE 1-BUTANOL DEHYDROGENASE BDH.
 OS Pseudomonas butanovora.
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 OC Pseudomonas.
 OX NCBI_TaxID-86174;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Vargnal A.S., Arp D.J., Sayavedra-Soto L.A.;
 RT Characterization of the expression of two distinct alcohol
 RT dehydrogenases involved in butane metabolism in Pseudomonas
 RT butanovora.
 RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF355798; AAK27220.2;
 SO SEQUENCE 691 AA; 75070 MW; 4FC7PD20CDA14E64 CRC64;

Query Match 16.9%; Score 462.5; DB 2; Length 691;
 Best Local Similarity 25.9%; Pred. No. 2.8e-25;
 Matches 145; Conservative 84; Mismatches 203; Indels 127; Gaps 20;

QY 39 GEMISYGNQENYRHSPLTQITTEVGOLOLWARGMQPK-VQVTPLIHGVMTLANPG 97
 Db 41 GEMRTHGYDAGTRYSPLAQITTPNAKELGLWMSYDESSRGVEATPIVVDGMYTATW 100
 QY 98 DYIOADAKTGDILMEHRRQIPNTATINSFGE-PTKMAIVANGYVAGSTQYSPFCGE 156
 Db 101 SYVAALDVRSKRRLMTYDPEVPREKKNACDDVNRVA-VHEGKVFVS----- 149
 QY 157 VSGH-----DSATGEELW-RNYFI-----PRA--GEEDDETGMNDYEAAMTGA 198
 Db 150 LDGRVLADARTGRKWRBNTLDDDKPYITIGAPRVKQVYVIGGAEFGVGTITAY 209
 QY 199 -----GQI-----TYDPVNTLVHVGSTAVG----- 218
 Db 210 DPTAASRGVVGPGGDSLPFEDASMEAAKTDWDPAGOVYLSGRRRHGVELDGLYRKAGF 269
 QY 219 -----PASETORGTPGGLTYGNTREFAVRPTGEIYWRHQTLPROMWDECFEM 268
 Db 270 CCTSAPATPSPWSHRKRSPAGGDNLTASTIVALRPTGVEYVHYQOTPADNMDYSTODL 329
 QY 269 MYTVADV--OPSTMEGLSINPNAATGERRVLTGPCKTGTWMOFDAETGEFLMARDTN 326
 Db 330 ILADIELGRP-----RKVILHAP-KNGFEFVIDRDKGFIQAQNFV 370
 QY 327 YONMIESIDENGIVTVN-EDAILKELDEYDVCPTFLGDRMPAALNDPSGIFYIFLPLNN 385

```

Db 371 PNNMATGIDENRPIENEGAMPGLSRF---PAPARFNMHMSVSPQGLAYFAQ 427
OY 386 VCYDMAVADDEFTSMDEVNTSN-----VTKLPECKMIGRIDALDISTR 430
Db 428 IFL-VLQEDKNMS---YNOAPGQAMAGIOMNLMLNPPAPSOQFGRILMDPVQOK 482
OY 431 TMSVERAANSPVLSGGVLEFGDRFRFALSOETGFEMLQTLATVASGOATSYE 490
Db 483 EWRREHSPWNGSTLVAGNVFQGTADARLFLDARDGKELMSAPMGIVTAPVTE 542
OY 491 VDMQVVAIA-GGGVSYGS 508
Db 543 VDGKQVSLAVGWGVYGN 561

RESULT 11
024759 PRELIMINARY; PRT; 633 AA.
ID 024759;
AC 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE METANOL DEHYDROGENASE ALPHA SUBUNIT PRECURSOR (EC 1.1.99.8).
GN KXAF.
OS Hypomicrobium methylovorum.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Hypomicrobium group; Hypomicrobium.
ON NCBI_TaxID=84;
RN [1]
RC SEQUENCE FROM N.A.
RX MEDLINE-97457202; PubMed-93111140;
RA Tanaka Y., Yoshida T., Watanabe K., Izumi Y., Mitsuura T.;
RT "Cloning and analysis of methanol oxidation genes in the methylotroph
RT Hypomicrobium methylovorum GM2."
RL FEMS Microbiol. Lett. 154:397-401(1997).
DR EMBL; AB004097; BAA23272.1; -.
DR HSSP; P38539; AAAH.
DR InterPro; IPR001479; Bac_P00.
DR InterPro; IPR002372; Bac_P00_repeat.
DR Pfam; PF01011; Bacterial_P00_7.
DR PROSITE; PS00363; BACTERIAL_P00_1; 1.
DR PROSITE; PS00364; BACTERIAL_P00_2; 1.
KW Signal; Oxidoreductase.
FT SIGNAL 1 34
FT CHAIN 35 633 POTENTIAL.
FT SEQUENCE 633 AA; 69853 MW; B47A23A2779E2C1B CRC64;
SQ

Query Match 15.6%; Score 428.5; DB 2; Length 633;
Best Local Similarity 24.6%; Pred. No. 7,4e-23;
Matches 146; Conservative 98; Mismatches 235; Indels 115; Gaps 22;

```

```

OY 228 PGTLTGNTREFAVRPDITGELVWRHQTLPDNDMDQETFEEMATVNDVORSTEMEGLQST 287
Db 305 PGDNKW-TMTITARDADGKMKFGYOKTPHDEMDFAGVNYIMSE-----QDKE----- 353
OY 288 NPNATGERRVILVPCPKTGTMMQFDATGFEFLA---RPTNOMNIESIDEINGIVTNE 344
Db 354 -----GKRRLKHPDNKGIYTLDRNGDLISAKLDIT--VNRKHDKLSGLPYRD 405
OY 345 DALIELDVE-YDVCPFLGGRDWPBALNPDGSIYFPIPLNWCYDMAVADDEFTSMDEV 403
Db 406 PEFGRMDKGTGLCPASMGYHNOGHOSYDPTKQLFPMGINHICMDPEPMLPYRAGOFF 465
OY 404 NTSNVTCLP-PEKDM-----IGRIDALDISTRILMSVERAANSPVLSGGVLENGG 457
Db 466 VGATLMYGPGRGDRQNYLIGLOIKAYNAINTYKWEHMERFSWGGTLTATGMLVYGT 525
OY 458 TDRYFRALSOETGFEMLQTLATVASGOATSYEVDQMOYVIAAG-GVSYGSL 510
Db 526 LDGFLKARSDPGELMKRLPSGVIGTPMTEHKQVYIIVNMGVGVGVL 579

RESULT 12
09AGW3 PRELIMINARY; PRT; 623 AA.
ID 09AGW3;
AC 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE 1-BUTANOL DEHYDROGENASE BOH PRECURSOR.
OS Pseudomonas butanovora.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
ON NCBI_TaxID=86174;
RN [1]
RC SEQUENCE FROM N.A.
RA Sayavedra-Soto L.A., Vangnal A.S., Arp D.J.;
RT "Characterization of the expression of two distinct alcohol
RT dehydrogenases involved in butane metabolism in Pseudomonas
RT butanovora."
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF326086; AAK15506.1; -.
DR HSSP; 092437; 1FLG.
DR InterPro; IPR002372; Bac_P00_repeat.
DR Pfam; PF01011; Bacterial_P00_4.
KW Signal.
FT SIGNAL 1 29
FT SEQUENCE 623 AA; 67553 MW; A573A1C64AC12D55 CRC64;
SQ

Query Match 15.5%; Score 424; DB 2; Length 623;
Best Local Similarity 25.2%; Pred. No. 1.5e-22;
Matches 159; Conservative 87; Mismatches 218; Indels 168; Gaps 26;

```


QY	208	NLVHNGTAAGVACSTAGTGGTGG-----TLGTYNDREAVRBDTEIYWRHDTLPDWMQD	262
Db	249	NLFYTGSSNPAPMYMTM--FDGKKWTMTINGRD-----DTMAKKGYOKTQPHDMDF	300
QY	263	ECTFEAMYTNDVDPSTBMEGLSINPNATGERRVLVGPCKTGTGMQDAETGEFLMA	322
Db	301	AGVNMVLTLD-----QEVN-----GKMTPLSHIDRNGILYTLNRENGNIIYA	343
QY	323	BDTN-YQMIIESIDENGIVTNEAIIKELDVE-YDVCPPFLGGRDMPMAALNDGSIYF	360
Db	344	EKVDPAVNVFKKVDLKTGTPTPRDEFAFRMDHKNTICPSAMGHNGQSYDPESFTLY	403
QY	361	IPLNWVCYDMAVVOEFTSMVDYVTSVNTKLP-----PGKMLGRIDAIDISTGRITMSYE	436
Db	404	AGLNLHICDMEPEMLPYRAGQFEFGATILAMYPGENGPTEKKEMOIRAFDILTTCRAKWTKW	463
QY	437	RAANYSFVLSTGGGVFENGSTDRYFRALSOETETLMOTRLATVAAQSALIEVQMOY	496
Db	464	EKFAAMGGTLYTKGVLVWYATLDQYLEKALDNKQSKELMNNKMPSSGGISGPMYTSFPGKQY	523
QY	497	VAIAGGVSYSYG 509	
Db	524	I-----GSMITGV 531	

Query Match	14.3%	Score 392.5	DB 2	Length 573
Best Local Similarity	24.6%	Pred. No. 2.7e-20		
Matches	136	Conservative	84	Mismatches 218; Indels 115; Gaps 22.
QY	38	AGEWISYGNQENGRHSPLLQTITTEWGOQLWV-ARGMPKGVYQVPLPIIHGWYTL--	93	
DB	13	AGAMPATGAGYSSOHNSPPLAQINKSNVAKAMSFSFGLNCH-EGAPLYIDMMKYVS	71	
QY	94	ANGADVIOALDAKTDGDLMEHR-RQLPRLIALINSFGPTGMA-----IV--ANGVIYA	144	
DB	72	AFPNNTYALNLDPGKIVYQHRPKDOSTKAMCCDYYVDGLAYGAGQIYKROANSHLLA	131	
QY	145	-----GSTCYSP-----GC-----FVSGHDSATGEEL	168	
DB	132	LDAKTGKINMEVEVCDPKVSGSTIIPQAPVAKDYIWMSCSAELGVGRGAVNAAPDKTKEGLK	191	
QY	169	WR-----NYFIAPAGE--BEDETWGDYTFARMNTGA--WQIITYDPVT	207	
DB	192	WRPFATGSDSDVRLAKDENSANPHYGQGLGTIKWED--AAKIGGTNWGWYADPGL	248	

Fri May 24 11:27:34 2002

walick-934-135.pep.rspt

Page 10

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2002, 10:01:01 ; Search time 117.28 Seconds
(without alignments)
89.973 Million cell updates/sec

Title: US-08-934-506a-5_COPY_1_95

Perfect score: 498
Sequence: 1 MKPTSLWASGALALALAP.....OPKQVYTPILHDGYMTLAN 95

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A.GeneSeq_032802:*

- 1: /SIDSL/gcgdata/hold-geneSeq/geneSeq-emb1/AA1980.DAT:*
- 2: /SIDSL/gcgdata/hold-geneSeq/geneSeq-emb1/AA1981.DAT:*
- 3: /SIDSL/gcgdata/hold-geneSeq/geneSeq-emb1/AA1982.DAT:*
- 4: /SIDSL/gcgdata/hold-geneSeq/geneSeq-emb1/AA1983.DAT:*
- 5: /SIDSL/gcgdata/hold-geneSeq/geneSeq-emb1/AA1984.DAT:*
- 6: /SIDSL/gcgdata/hold-geneSeq/geneSeq-emb1/AA1985.DAT:*
- 7: /SIDSL/gcgdata/hold-geneSeq/geneSeq-emb1/AA1986.DAT:*
- 8: /SIDSL/gcgdata/hold-geneSeq/geneSeq-emb1/AA1987.DAT:*
- 9: /SIDSL/gcgdata/hold-geneSeq/geneSeq-emb1/AA1988.DAT:*
- 10: /SIDSL/gcgdata/hold-geneSeq/geneSeq-emb1/AA1989.DAT:*
- 11: /SIDSL/gcgdata/hold-geneSeq/geneSeq-emb1/AA1990.DAT:*
- 12: /SIDSL/gcgdata/hold-geneSeq/geneSeq-emb1/AA1991.DAT:*
- 13: /SIDSL/gcgdata/hold-geneSeq/geneSeq-emb1/AA1992.DAT:*
- 14: /SIDSL/gcgdata/hold-geneSeq/geneSeq-emb1/AA1993.DAT:*
- 15: /SIDSL/gcgdata/hold-geneSeq/geneSeq-emb1/AA1994.DAT:*
- 16: /SIDSL/gcgdata/hold-geneSeq/geneSeq-emb1/AA1995.DAT:*
- 17: /SIDSL/gcgdata/hold-geneSeq/geneSeq-emb1/AA1996.DAT:*
- 18: /SIDSL/gcgdata/hold-geneSeq/geneSeq-emb1/AA1997.DAT:*
- 19: /SIDSL/gcgdata/hold-geneSeq/geneSeq-emb1/AA1998.DAT:*
- 20: /SIDSL/gcgdata/hold-geneSeq/geneSeq-emb1/AA1999.DAT:*
- 21: /SIDSL/gcgdata/hold-geneSeq/geneSeq-emb1/AA2000.DAT:*
- 22: /SIDSL/gcgdata/hold-geneSeq/geneSeq-emb1/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	498	100.0	579	19	AAW37873
2	428	85.9	579	19	AAW37874
3	404	81.1	579	19	AAW37876
4	381.5	76.6	578	21	AAW37875
5	142	28.5	754	21	AAW35987
6	131.5	26.4	738	12	AAW31993
7	131.5	26.4	738	13	AAW20192
8	127.5	25.6	443	22	ABG24430
9	126.5	25.4	742	11	AAW05235
10	122.5	24.6	948	22	ABG25909
11	122.5	24.6	1510	22	ABG21573

12	120	24.1	740	20	AAW95019
13	73.5	14.8	1229	22	ABW71736
14	72.5	14.6	343	19	AAW35001
15	71	14.3	526	22	AAW98995
16	69	13.9	711	22	ABW69743
17	66	13.3	227	22	AAW56221
18	66	13.3	1419	22	ABW59208
19	65.5	13.2	289	21	AAW42562
20	65.5	13.2	1004	22	AAW90456
21	65	13.1	43	22	AAW59378
22	65	13.1	43	22	AAW71934
23	65	13.1	824	18	AAW09614
24	64.5	13.0	255	16	AAW92829
25	64.5	13.0	1171	22	ABW64387
26	64	12.9	229	22	AAW50083
27	64	12.9	238	22	AAW90966
28	64	12.9	371	22	AAW91604
29	64	12.9	487	22	AAW50084
30	64	12.9	1016	22	AAW34349
31	63.5	12.8	93	22	AAW01744
32	63.5	12.8	115	22	AAW01694
33	63	12.7	539	22	AAW64896
34	62.5	12.6	138	22	ABG14809
35	62.5	12.6	420	16	AAW91068
36	62.5	12.6	420	18	AAW07593
37	62.5	12.6	420	21	AAW90454
38	62.5	12.6	583	20	AAW55919
39	62.5	12.6	712	22	ABW70600
40	62	12.4	577	22	ABG13760
41	61.5	12.3	73	22	ABW17647
42	61.5	12.3	221	21	AAW48548
43	61.5	12.3	1669	22	ABW64003
44	61	12.2	553	15	AAW56966
45	61	12.2	553	15	AAW56967

ALIGNMENTS

RESULT 1
AAW37873
ID AAW37873 standard; Protein: 579 AA.

AC AAW37873;

DT 10-AUG-1998 (first entry)

DE Alcohol and/or aldehyde dehydrogenase A amino acid sequence.

KW Alcohol/aldehyde dehydrogenase A enzyme; recombinant organism;

KW aldehyde; ketone; carboxylic acid; L-sorbose; D-sorbitol;

KW 2-Keto-L-gulonic acid; L-ascorbic; inhibition.

OS Gluconobacter oxydans.

FT Key

FT Peptide

FT Protein

PN EP832974-A2.

PD 01-APR-1998.

PF 11-SEP-1997; 97EP-0115801.

PR 19-SEP-1996; 96EP-0115001.

PA (HOFF) HOFFMANN LA ROCHE & CO AG F.

PI Asakura A, Hoshino T, Ojima S, Shinjoh M, Tomiyama N;

Sorbitol dehydroge
Drosophila melanog
Bacillus thuringien
Xanthomonas sp. ch
Drosophila melanog
Propionibacterium
Drosophila melanog
Human ORF2326
C glutamicum prote
Human brain expres
Human bone marrow
Pathogenic Staphyl
C179FV-PP variable
Drosophila melanog
Rat IgG1 Fc domain
C glutamicum prote
C glutamicum prote
TMFR-Fc fusion pro
Staphylococcus aur
Human gene 1 encod
Drosophila melanog
Novel human diagn
bEGF-linker-sapori
Basic fibroblast g
bEGF/cathepsin B s
Pseudomonas fluores
Drosophila melanog
Novel human diagn
Human nervous syst
Cochliobolus cardio
Drosophila melanog
CC49 VL-L-VH-L-VL-
CC49 VL-L-VH-L-VH-

DR WPI: 1998-195228/18.
 DR N-PSDB: AAV29051.
 XX
 PT Recombinant Gluconobacter oxydans alcohol and/or aldehyde
 PT dehydrogenase enzyme(s) - useful for converting substrate(s), e.g.
 PT L-sorbose or D-sorbitol to 2-keto-L-gulonic acid
 XX
 PS Claim 1; Pages 35-37; 59pp; English.
 CC
 CC This is the amino acid sequence for the Gluconobacter oxydans
 CC alcohol and/or aldehyde dehydrogenase A enzyme. The enzymes
 CC or recombinant organisms can be used to convert suitable substrates
 CC to aldehydes, ketones or carboxylic acids, especially to convert
 CC L-sorbose or D-sorbitol to 2-keto-L-gulonic acid, which can be
 CC converted to L-ascorbic acid by standard procedures. The derivatives
 CC of ADH enzymes have desired substrate specificity, higher affinity
 CC to a substrate, lower affinity to an inhibitory compound, higher
 CC stability against temperature and/or pH and higher catalytic speed.
 CC
 SQ Sequence 579 AA;
 Query Match 100.0%; Score 498; DB 19; Length 579;
 Best Local Similarity 100.0%; Pred. No. 8e-49;
 Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 MKPTSLMASAGALALIAFAFAQVTPYTDIELANPAGEWISYSGONENYRHSPLTQIT 60
 DB 1 mkptslmasagallalaiafaafagvtpvtdellannpagewisysgngenyhspltglt 60
 OY 61 TENVGQLQIYWARGMQPKVQVTPPLIHGVMYLAN 95
 DB 61 tenvgqlqiywargmqpkvqvtpplihdgvmylan 95
 RESULT 2
 AAW37874
 ID AAW37874 standard; Protein; 579 AA.
 AC AAW37874;
 XX
 DT 10-AUG-1998 (first entry)
 DE Alcohol and/or aldehyde dehydrogenase A' amino acid sequence.
 XX
 KW Alcohol/aldehyde dehydrogenase A' enzyme; recombinant organism;
 KW aldehyde; ketone; carboxylic acid; L-sorbose; D-sorbitol;
 KW 2-keto-L-gulonic acid; L-ascorbic; inhibition.
 XX
 OS Gluconobacter oxydans.
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..23
 FT Protein /note= "signal peptide"
 FT 24..579
 FT /note= "mature protein"
 XX
 EP832974-A2.
 XX
 PN 01-APR-1998.
 PD
 XX
 PE 11-SEP-1997; 97EP-0115801.
 PR 19-SEP-1996; 96EP-0115001.
 XX
 PA (HOFF) HOFFMANN LA ROCHE & CO AG F.
 PI Asakura A, Hoshino T, Ojima S, Shinjoh M, Tomiyama N;
 DR WPI: 1998-195228/18.
 DR N-PSDB: AAV29052.
 XX
 PT Recombinant Gluconobacter oxydans alcohol and/or aldehyde

PT dehydrogenase enzyme(s) - useful for converting substrate(s), e.g.
 PT L-sorbose or D-sorbitol to 2-keto-L-gulonic acid
 XX
 PS Claim 1; Pages 38-40; 59pp; English.
 CC
 CC This is the amino acid sequence for the Gluconobacter oxydans
 CC alcohol and/or aldehyde dehydrogenase A' enzyme. The enzymes
 CC or recombinant organisms can be used to convert suitable substrates
 CC to aldehydes, ketones or carboxylic acids, especially to convert
 CC L-sorbose or D-sorbitol to 2-keto-L-gulonic acid, which can be
 CC converted to L-ascorbic acid by standard procedures. The derivatives
 CC of ADH enzymes have desired substrate specificity, higher affinity
 CC to a substrate, lower affinity to an inhibitory compound, higher
 CC stability against temperature and/or pH and higher catalytic speed.
 CC
 SQ Sequence 579 AA;
 Query Match 85.9%; Score 428; DB 19; Length 579;
 Best Local Similarity 88.4%; Pred. No. 9.5e-41;
 Matches 84; Conservative 3; Mismatches 8; Indels 0; Gaps 0;
 OY 1 MKPTSLMASAGALALIAFAFAQVTPYTDIELANPAGEWISYSGONENYRHSPLTQIT 60
 DB 1 mkptslmasagallalaiafaafagvtpvtdellannpagewisysgngenyhspltglt 60
 OY 61 TENVGQLQIYWARGMQPKVQVTPPLIHGVMYLAN 95
 DB 61 tenvgqlqiywargmqpkvqvtpplihdgvmylan 95
 RESULT 3
 AAW37876
 ID AAW37876 standard; Protein; 579 AA.
 AC AAW37876;
 XX
 DT 10-AUG-1998 (first entry)
 DE Alcohol and/or aldehyde dehydrogenase B amino acid sequence.
 XX
 KW Alcohol/aldehyde dehydrogenase B enzyme; recombinant organism;
 KW aldehyde; ketone; carboxylic acid; L-sorbose; D-sorbitol;
 KW 2-keto-L-gulonic acid; L-ascorbic; inhibition.
 XX
 OS Gluconobacter oxydans.
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..23
 FT Protein /note= "signal peptide"
 FT 24..579
 FT /note= "mature protein"
 XX
 EP832974-A2.
 XX
 PN 01-APR-1998.
 PD
 XX
 PE 11-SEP-1997; 97EP-0115801.
 PR 19-SEP-1996; 96EP-0115001.
 XX
 PA (HOFF) HOFFMANN LA ROCHE & CO AG F.
 PI Asakura A, Hoshino T, Ojima S, Shinjoh M, Tomiyama N;
 DR WPI: 1998-195228/18.
 DR N-PSDB: AAV29054.
 XX
 PT Recombinant Gluconobacter oxydans alcohol and/or aldehyde
 PT dehydrogenase enzyme(s) - useful for converting substrate(s), e.g.
 PT L-sorbose or D-sorbitol to 2-keto-L-gulonic acid
 XX
 PS Claim 1; Pages 44-46; 59pp; English.

CC In sequences AAC83156 and AAC83157.
XX
SQ Sequence 754 AA;

Query Match 28.58; Score 142; DB 21; Length 754;
Best Local Similarity 40.28; Pred. No. 1.3e-07;
Matches 35; Conservative 10; Mismatches 34; Indels 8; Gaps 2;

OY 14 LALAAPAFVAVPTVDE-----TDELLANPAGEMISYCONQENYRHSPLTQTITTEVQ 66
DB 17 lgcasaalafcatcpvalaactgtatitnaqdhpgdmsygrtysqgrysplditkdnasn 76
OY 67 LQVWANGMPGKQVQV-TPLIHGVMY 92
DB 77 lklawhylditnrgqegtpivdgymy 103

RESULT 6

AAR13993
ID AAR13993 standard; Protein; 738 AA.

XX AAR13993;

XX 09-DEC-1991 (first entry)

DE A.altoacetigenes membrane-bound ADH 72KD sub-unit.

XX alcohol dehydrogenase complex; carboxylic acid production.

OS Acetobacter altoacetigenes.

XX EP448969-A.

XX 02-OCT-1991.

XX 26-FEB-1991; 91EP-0102793.

XX 26-MAR-1990; 90JP-0073440.

PR 26-FEB-1990; 90JP-0042391.

XX (NAKA-) NAKANO VINEGAR KK.

PI Tamaki T, Takemura H, Tayama K, Fukaya M, Okumura H;

PI Kawamura Y;

DR WPI: 1991-289462/40.
DR N-PSDB: AAQ13580.

PT Gene for membrane-bound alcohol dehydrogenase complex - obt'd.
PT from Acetobacter altoacetigenes, used for prodn. of enzyme for
PT converting alcohol to acid

PS Disclosure; Fig 3; 36pp; English.

CC Total DNA was prepared from A.altoacetigenes MH-24, digested with
CC PstI and SmaI and ligated to PstI-SmaI cleaved pUC18. The ligation
CC mixture was used to transform E.coli JM109. Probes were designed
CC based on the N-terminal amino acid sequence of the ADH complex
CC isolated from A. altoacetigenes (see AAQ13582-Q13584). The gene
CC encoding the 72KD ADH subunit was isolated and sequenced. The
CC directly sequenced N-terminal region of the purified 72KD sub-unit
CC corresponds to the sequence beginning at residue 36 of the deduced
CC sequence. This suggests that the first 35 N-terminal amino acids
CC form a leader peptide involved in secretion of the mature 72KD
CC protein. The deduced amino acid sequence has 77 per cent homology
CC with the same enzyme from A. aceti K6033. See also AAQ13581.

SQ Sequence 738 AA;

Query Match 26.4%; Score 131.5; DB 12; Length 738;
Best Local Similarity 32.6%; Pred. No. 2e-06;

Matches 30; Conservative 15; Mismatches 38; Indels 9; Gaps 2;

OY 10 SAGALALAAPAFVAVPTVDE-----LLANPAGEMISYCONQENYRHSPLTQTIT 61
DB 17 tagtcaalissyatacnaadddggatgaatlnadhpnmwlytrtysdgryspldqtnr 76
OY 62 ENVGQLQVWANGMPGKQVQV-TPLIHGVMY 92
DB 77 snvgnlklawylditnrgqegtpivdgymy 108

RESULT 7

AAR20192
ID AAR20192 standard; Protein; 738 AA.

XX AAR20192;

XX 16-APR-1992 (first entry)

DE ADH complex protein (mol.wt. 72.000).

XX Alcohol dehydrogenase; acetic acid; fermentation.

OS Acetobacter altoacetigenes NH-24.

XX JP0326988-A.

XX 27-NOV-1991.

XX 26-MAR-1990; 90JP-0073440.

XX 26-FEB-1990; 90JP-0042301.

PR 26-MAR-1990; 90JP-0073440.

XX (NAKA-) NAKANO SUTEN KK.

XX WPI: 1992-019325/03.

DR N-PSDB: AAQ20383.

PT Alcohol dehydrogenase complex structural gene - used in plasmid
PT and enhancing efficiency of acetic acid fermentation for
PT transformed acetic acid bacteria

PS Disclosure; Fig 3(1-3); 21pp; Japanese.

CC Acetobacter transformed with the sequence encoding this protein can
CC enhance the efficiency of acetic acid fermentation. The ADH complex
CC can be easily extracted from the bacteria and purified and it can be
CC used for the determination of an alcohol.
CC See also AAQ20383-84, and -86-88.

SQ Sequence 738 AA;

Query Match 26.4%; Score 131.5; DB 13; Length 738;
Best Local Similarity 32.6%; Pred. No. 2e-06;
Matches 30; Conservative 15; Mismatches 38; Indels 9; Gaps 2;

OY 10 SAGALALAAPAFVAVPTVDE-----LLANPAGEMISYCONQENYRHSPLTQTIT 61
DB 17 tagtcaalissyatacnaadddggatgaatlnadhpnmwlytrtysdgryspldqtnr 76
OY 62 ENVGQLQVWANGMPGKQVQV-TPLIHGVMY 92
DB 77 snvgnlklawylditnrgqegtpivdgymy 108

RESULT 8

ABG24430
ID ABG24430 standard; Protein; 443 AA.

XX ABG24430;

DT 18-FEB-2002 (first entry)

XX Novel human diagnostic protein #24421.

DE Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW Food supplement; medical imaging; diagnostic; genetic disorder.

XX Homo sapiens.

XX WO200175067-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US08631.

XX 31-MAR-2000; 2000US-0540217.

XX 23-AUG-2000; 2000US-0649167.

XX (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.

XX N-PSDB; AAS88617.

XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity

PS Claim 20; SEQ ID No 54789; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and
XX polypeptide (II) sequences. (I) is useful as hybridization probes,
XX polymerase chain reaction (PCR) primers, oligomers, and for chromosome
XX and gene mapping, and in recombinant production of (II). The
XX polynucleotides are also used in diagnostics as expressed sequence tags
XX for identifying expressed genes. (I) is useful in gene therapy techniques
XX to restore normal activity of (II) or to treat disease states involving
XX (II). (II) is useful for generating antibodies against it, detecting or
XX quantitating a polypeptide in tissue, as molecular weight markers and as
XX a food supplement. (II) and its binding partners are useful in medical
XX imaging of sites expressing (II). (I) and (II) are useful for treating
XX disorders involving aberrant protein expression or biological activity.
XX The polypeptide and polynucleotide sequences have applications in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits to assess biodiversity
XX and to produce other types of data and products dependent on DNA and
XX amino acid sequences. ABG00010-ABG30377 represent novel human
XX diagnostic amino acid sequences of the invention.
XX Note: The sequence data for this patent did not appear in the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 443 AA:

Query Match 25.6%; Score 127.5; DB 22; Length 443;
Best Local Similarity 33.7%; Pred. No. 3.1e-06;
Matches 35; Conservative 15; Mismatches 29; Indels 25; Gaps 5;

OY 6 LLMAS-----AGALLAAPAFQVTPYTDLLANPRAGWISYSGONEWRHSPLTQ 58
DB 168 ltwagfndpgeingcllsadelpaea-ispvadt-----dwpaygrnqegqfslp 218
OY 59 ITTENWGLOLVMA-----RGMOPGV--QVTPLLIHGVMYL 93
DB 219 inadhvnhlkeawfrtgdvkvgnpdgeitnevrplkvgtllyl 262

RESULT 9
AAR05235
ID AAR05235 standard; protein: 742 AA.

XX AAR05235:

XX 04-AUG-1990 (first entry)

DE Amino acid sequence of alcohol dehydrogenase (ADH) encoded by bases
DE 1-2229 of alcohol dehydrogenase (ADH) gene.

XX Alcohol dehydrogenase (ADH) gene; Acetobacter pasteurianus IFO 3191;
KW Acetobacter aceti K1006 (FERM-7528); plasmid PA721.

XX Acetobacter aceti strain K1006 (FERM-7528).

XX JP02000452-A.

XX 05-JAN-1990.

XX 30-OCT-1987; 87JP-0273190.

XX 30-OCT-1987; 87JP-0273190, JP-075069.

XX (QPP) QP CORP (QPP-).

XX WPI; 1990-047990/07.

XX N-PSDB; AAQ91811.

XX Cloning DNA, plasmid and microbe contg. it
PT contains alcohol dehydrogenase gene derived from Acetobacter
PT aceti K1006 strain

XX Disclosure; Fig 4; 8pp; Japanese.

PS Also new are a recombinant plasmid contg. its encoding DNA, and a
XX microorganism transformed with the plasmid.

XX A DNA fragment was obtd. from A. aceti
XX strain K1006 and was transferred to an ADH-defective strain. The
XX resulting plasmid, PA721, contg. its gene was inserted into A.
XX pasteurianus strain IFO 3191 by the triparent method. Its gene is
XX useful for improving Acetobacter culturing.

XX Sequence 742 AA:

Query Match 25.4%; Score 126.5; DB 11; Length 742;
Best Local Similarity 38.8%; Pred. No. 7.7e-06;
Matches 33; Conservative 11; Mismatches 38; Indels 3; Gaps 2;

OY 11 AGALLAAPAFQVTPYTDLL--ANPRAGWISYSGONEWRHSPLTQITENWGLO 68
DB 23 aaalpyaavparaagqgnlgealhhaddhpenwisgyrtysqfyspldlnrnvgdlk 82
OY 69 LVWARGMOPGVQ--VTPLLIHGVMY 92
DB 83 llyyrltdnrggeatpplvvdgimy 107

RESULT 10

ABG25909
ID ABG25909 standard; protein: 948 AA.

XX ABG25909;

XX 18-FEB-2002 (first entry)

XX Novel human diagnostic protein #25900.

DE Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW Food supplement; medical imaging; diagnostic; genetic disorder.

XX Homo sapiens.

XX WO200175067-A2.

XX

PD 11-OCT-2001.
 XX 30-MAR-2001; 2001WO-US08631.
 XX 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 XX (HYSE-) HYSEQ INC.
 XX Dermanac RT, Liu C, Tang YT;
 PI WPI: 2001-639362/73.
 DR N-PSDB; AAS90096.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity
 XX
 PS Claim 20; SEQ ID No 56268; 103bp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SO Sequence 948 AA;
 Query Match 24.6%; Score 122.5; DB 22; Length 948;
 Best Local Similarity 34.4%; Pred. No. 3.1e-05;
 Matches 33; Conservative 14; Mismatches 24; Indels 25; Gaps 5;
 QY 6 LLMAS-----AGATALLAARFAQVTPVTDLLANPPAGEMISYGQENYRHSPLTQ 58
 Db 310 ltwagfndpgeingtlisadatpaea-ispvadq-----dwpaygrngqgrtspklq 360
 QY 59 ITTENVGOLQLVWA-----RGMOPGKV--QVTP 85
 Db 361 inadvnmlkeawvrtgtvdkpndpgeltnevtpl 396
 RESULT 11
 ABG21573
 ID ABG21573 standard; Protein; 1510 AA.
 XX
 AC ABG21573;
 XX
 DT 18-FEB-2002 (first entry)
 XX
 DE Novel human diagnostic protein #21564.
 XX
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 XX food supplement; medical imaging; diagnostic; genetic disorder.
 OS Homo sapiens.

XX
 FN WO200175067-A2.
 XX 11-OCT-2001.
 PD 30-MAR-2001; 2001WO-US08631.
 XX 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 XX (HYSE-) HYSEQ INC.
 XX Dermanac RT, Liu C, Tang YT;
 PI WPI: 2001-639362/73.
 DR N-PSDB; AAS85760.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity
 XX
 PS Claim 20; SEQ ID No 51932; 103bp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SO Sequence 1510 AA;
 Query Match 24.6%; Score 122.5; DB 22; Length 1510;
 Best Local Similarity 34.4%; Pred. No. 5.6e-05;
 Matches 33; Conservative 14; Mismatches 24; Indels 25; Gaps 5;
 QY 6 LLMAS-----AGATALLAARFAQVTPVTDLLANPPAGEMISYGQENYRHSPLTQ 58
 Db 310 ltwagfndpgeingtlisadatpaea-ispvadq-----dwpaygrngqgrtspklq 360
 QY 59 ITTENVGOLQLVWA-----RGMOPGKV--QVTP 85
 Db 361 inadvnmlkeawvrtgtvdkpndpgeltnevtpl 396
 RESULT 12
 AAW95019
 ID AAW95019 standard; Protein; 740 AA.
 XX
 AC AAW95019;
 XX
 DT 21-MAY-1999 (first entry)
 XX
 DE Sorbitol dehydrogenase (SUDH) protein sequence.
 XX
 KW Sorbitol dehydrogenase; SUDH; open reading frame; ORF2; L-sorbitose;

[illegible]

PI	Lam DE, Mathur EJ;
XX	
DR	WPI; 1998-018435/02.
DR	N-PSDB; AAT94209.
XX	

XX Endoglucanase(s), preferably form archaeal bacterium, ABEPI 1a -
PT useful to degrade carboxymethylcellulose and hydrolyse of
PT beta-1,4-glycosidic bonds in cellulose

PS Claim 1; Fig 1Q; 164pp; English.
 XY

CC virus protein comprises an endoglucanase of *Bacillus thermoleovorans*
CC (clone 68GC1) that is capable of degrading carboxymethylcellulose
CC and of hydrolysing the beta-1,4-glycosidic bonds in cellulose. It
CC has homology to an endoglucanase of *Arthrobacterium* sp. (see
CC *AMW31985*). It can be produced from native cells or from recombinant
CC host cells, especially prokaryotic host cells transformed with a
CC plasmid or virus-derived vector including the endoglucanase DNA
CC (see *AMW4209*). 24 Endoglucanases (see *AMW31986-W5008*) are claimed
CC biomass into fuels and chemicals; for use in the conversion of plant
CC animal feed, waste treatment, and in the fruit juice and brewing
CC industries for the clarification and extraction of juices.

Sequence 343 AA;

Query Match	14.6%	Score 72.5;	DB 19;	Length 343;
Best Local Similarity	35.3%	Pred. No. 4.8;		
Matches 18; Conservative	6;	Mismatches 26;	Indels 1;	Gaps 1;

20 VIDELLANPAGEMISIGQNGENRHSPLTQTITTENQGLQVLWARGMOPG 78
 106 vlenasnplegewekgvtkwesfaldataffenhgkryyvwag-kdpg 155

RESULT 15

AA099895 standard; Protein; 526 AA.

AC AAB99895;

21-SEP-2001 (first entry)

..... SP: cholesteryl esterase protein SEQ ID NO:1
XX

xx xanthomonas; cholesterol esterase.

Xanthomonas sp.

PN WO200144451-A1.

21-JIN-2001

XX
XX
10
XXXX
00000

[illegible]

14 DEC-1999; 99JP-0353/95.

(KIKK) KIKKOMAN CORP.

Matsumoto K, Furukawa K, Koyama Y, Ichikawa Y, Katiyama N.

WPI: 2001-408478/A3

N-PSDB; AAH44284.

production of recombinant cholesterol esterase gelle of *Xanthomonas* origin for efficient

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

The present sequence represents a cholesterol esterase originating in *Xanthomonas* sp. No. 81-13. The cholesterol esterase can be used for the efficient production of recombinant cholesterol esterase by microbial

CC	culture.	526
XX		AA;
SQ	Sequence	

Query Match	14.38;	Score 71;	DB 22;	Length 526;
Best Local Similarity	40.98;	Pred. No. 13;		
Matches	18;	Conservative	3;	Mismatches 17;
				Indels 6;
				Gaps 1.

QY 8 WASAGALALLAPAFQAQVTPVTDELLNPPA-----GEWISYC 45

Db 95 wdsagllidiagaptvedvsavidwalantpdaipdaigagsgisysg 138

Search completed: May 24, 2002, 10:04:59
Job time: 238 sec

•
•
Fri May 24 11:27:21 2002

us-08-934-506a-5_copy_1_95.rag

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2002, 09:59:06 ; Search time 44.35 Seconds
(Without alignments)
52.321 Million cell updates/sec

Title: US-08-934-506a-5_COPY_1_95
Perfect score: 498
Sequence: 1 MKPTSLMASGALALAAP.....QPGKVOVTPLIHDGVTIAN 95

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08
Maximum Match 1008
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/prodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/prodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/prodata/2/1aa/6A.COMB.pep:*
4: /cgn2_6/prodata/2/1aa/6B.COMB.pep:*
5: /cgn2_6/prodata/2/1aa/PCRTUS.COMB.pep:*
6: /cgn2_6/prodata/2/1aa/Backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	142	28.5	754	4 US-09-296-284-4	Sequence 4, Appl
2	133	26.7	720	4 US-09-286-284-25	Sequence 25, Appl
3	131.5	26.4	738	1 US-07-985-458-3	Sequence 3, Appl
4	120	24.1	739	3 US-07-136-251-2	Sequence 2, Appl
5	61	12.2	275	3 US-08-463-903-8	Sequence 8, Appl
6	61	12.2	275	4 US-07-935-695-8	Sequence 8, Appl
7	61	12.2	280	3 US-08-463-903-10	Sequence 10, Appl
8	61	12.2	280	4 US-07-935-695-10	Sequence 10, Appl
9	61	12.2	282	3 US-08-463-903-12	Sequence 12, Appl
10	61	12.2	282	4 US-07-935-695-12	Sequence 12, Appl
11	61	12.2	553	2 US-08-263-911-7	Sequence 7, Appl
12	61	12.2	553	2 US-08-263-911-9	Sequence 9, Appl
13	60.5	12.1	587	2 US-08-436-664-23	Sequence 23, Appl
14	60.5	12.1	587	3 US-09-135-642-23	Sequence 23, Appl
15	60.5	12.1	587	3 US-08-394-232A-23	Sequence 23, Appl
16	60.5	12.1	587	3 US-08-394-232A-23	Sequence 23, Appl
17	60.5	12.1	588	5 US-08-436-664-25	Sequence 25, Appl
18	60.5	12.1	588	3 US-09-135-642-25	Sequence 25, Appl
19	60.5	12.1	588	3 US-08-394-232A-25	Sequence 25, Appl
20	60.5	12.1	588	5 PCT-US95-04080-25	Sequence 25, Appl
21	60.5	12.1	590	2 US-08-436-664-27	Sequence 27, Appl
22	60.5	12.1	590	3 US-09-135-642-27	Sequence 27, Appl
23	60.5	12.1	590	3 US-08-394-232A-27	Sequence 27, Appl
24	60.5	12.1	590	3 PCT-US95-04080-27	Sequence 27, Appl
25	60.5	12.1	624	3 US-09-057-969-3	Sequence 3, Appl
26	60.5	12.1	876	2 US-08-436-664-20	Sequence 20, Appl
27	60.5	12.1	876	2 US-08-436-664-32	Sequence 32, Appl

28	60.5	12.1	876	2 US-08-436-664-34	Sequence 34, Appl
29	60.5	12.1	876	3 US-09-135-642-20	Sequence 20, Appl
30	60.5	12.1	876	3 US-09-135-642-32	Sequence 32, Appl
31	60.5	12.1	876	3 US-09-135-642-34	Sequence 34, Appl
32	60.5	12.1	876	3 US-08-394-232A-20	Sequence 20, Appl
33	60.5	12.1	876	3 US-08-394-232A-32	Sequence 32, Appl
34	60.5	12.1	876	3 US-08-394-232A-34	Sequence 34, Appl
35	60.5	12.1	876	5 PCT-US95-04080-20	Sequence 20, Appl
36	60.5	12.1	876	5 PCT-US95-04080-32	Sequence 32, Appl
37	60.5	12.1	876	5 PCT-US95-04080-34	Sequence 34, Appl
38	60.5	12.1	954	3 US-09-057-969-2	Sequence 2, Appl
39	60	12.0	251	4 US-08-858-207A-354	Sequence 354, Appl
40	59.5	11.9	291	4 US-09-286-691-28	Sequence 28, Appl
41	59.5	11.9	291	4 US-09-687-147-28	Sequence 28, Appl
42	59	11.8	176	4 US-09-213-343-3	Sequence 3, Appl
43	59	11.8	176	4 US-09-413-814-85	Sequence 85, Appl
44	59	11.8	223	2 US-08-121-436A-2	Sequence 2, Appl
45	59	11.8	568	4 US-09-160-065-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-09-296-284-4
Sequence 4, Application US/09296284A
Patent No. 6204040
GENERAL INFORMATION:
APPLICANT: Choi, Eun-Sung
APPLICANT: Rhee, Sang-Ki
APPLICANT: Lee, Eun-Hae
TITLE OF INVENTION: Gluconobacter Suboxydans Sorbitol Dehydrogenase, Genes
FILE OF INVENTION: and Methods of Use Thereof
FILE REFERENCE: 1533.0870000
CURRENT APPLICATION NUMBER: US/09/296, 284A
CURRENT FILING DATE: 1999-04-22
NUMBER OF SEQ. ID NOS: 87
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 4
LENGTH: 754
TYPE: PRT
ORGANISM: Gluconobacter suboxydans
US-09-296-284-4

Query Match 28.5%; Score 142; DB 4; Length 754;
Best Local Similarity 40.2%; Pred. No. 9.2e-09;
Matches 35; Conservative 10; Mismatches 34; Indels 8; Gaps 2;
QY 14 LALLAAPAFAGVTPY-----TDELLANPPAGEMISYGNQENTRHSPLOTITTENVGQ 66
DB 17 LGCAGAAAFACATSPVALAEDGTATNADQHPGDMSYGRYSRQRYSPIDQITKNASN 76
QY 67 LQLVWARGNQPKQVY-TPLIHDGVMY 92
DB 77 LKLANHYDIDTNRGEGPPLIVDGMV 103
RESULT 2
US-09-296-284-25
Sequence 25, Application US/09296284A
Patent No. 6204040
GENERAL INFORMATION:
APPLICANT: Choi, Eun-Sung
APPLICANT: Rhee, Sang-Ki
APPLICANT: Lee, Eun-Hae
TITLE OF INVENTION: Gluconobacter Suboxydans Sorbitol Dehydrogenase, Genes
FILE OF INVENTION: and Methods of Use Thereof
FILE REFERENCE: 1533.0870000
CURRENT APPLICATION NUMBER: US/09/296, 284A
CURRENT FILING DATE: 1999-04-22
NUMBER OF SEQ. ID NOS: 87
SOFTWARE: PatentIn Ver. 2.0

OTHER INFORMATION: :
US-07-935-695-8

Query Match 12.2%; Score 61; DB 4; Length 275;
Best Local Similarity 25.0%; Pred. No. 13;
Matches 22; Conservative 14; Mismatches 24; Indels 28; Gaps 4;

Db 8 AAGLLLLAQAOPAMADIVMSQSPSSLPVSGEKVTLSCSSQSLYSGNKNLYAWQOK 67

QY 9 ASAGALALLAAPAFQVY-----PVT--DELLANPAGEWISYCONQENY----- 51

Db 8 AAGLLLLAQAOPAMADIVMSQSPSSLPVSGEKVTLSCSSQSLYSGNKNLYAWQOK 67

QY 52 -RHSPITQITTEENVGOLQVWARGMOPG 78

Db 68 PGQSP-----KLLIYMASARESG 85

RESULT 6
US-07-935-695-8
Sequence 8, Application US/07935695
GENERAL INFORMATION:
APPLICANT: Mezes, Peter S.
APPLICANT: Richard, Ruth A.
APPLICANT: Affholter, Joseph A.
APPLICANT: Kollie, Nicolas J.
TITLE OF INVENTION: Dimer and Multimer Forms of Single Chain Polypeptides
FILE REFERENCE: 40224A US
CURRENT APPLICATION NUMBER: US/07/935,695
CURRENT FILING DATE: 1992-08-21
PRIOR APPLICATION NUMBER: US 08/463,903
PRIOR FILING DATE: 1995-06-05
NUMBER OF SEQ ID NOS: 102
SOFTWARE: MS-word for Windows, Ver. 7.0
SEQ ID NO 8
LENGTH: 275
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
NAME/KEY: SCFY UHM8.1
LOCATION: 1..275
US-08-463-903-8

Query Match 12.2%; Score 61; DB 3; Length 275;
Best Local Similarity 25.0%; Pred. No. 13;
Matches 22; Conservative 14; Mismatches 24; Indels 28; Gaps 4;

Db 8 AAGLLLLAQAOPAMADIVMSQSPSSLPVSGEKVTLSCSSQSLYSGNKNLYAWQOK 67

QY 9 ASAGALALLAAPAFQVY-----PVT--DELLANPAGEWISYCONQENY----- 51

Db 8 AAGLLLLAQAOPAMADIVMSQSPSSLPVSGEKVTLSCSSQSLYSGNKNLYAWQOK 67

QY 52 -RHSPITQITTEENVGOLQVWARGMOPG 78

Db 68 PGQSP-----KLLIYMASARESG 85

RESULT 5
US-08-463-903-8
Sequence 8, Application US/08463903
GENERAL INFORMATION:
APPLICANT: Mezes, Peter S.
APPLICANT: Richard, Ruth A.
APPLICANT: Affholter, Joseph A.
APPLICANT: Kollie, Nicolas J.
TITLE OF INVENTION: Dimer and Multimer Forms of Single Chain Polypeptides
FILE REFERENCE: 40224A US
CURRENT APPLICATION NUMBER: US/08/463,903
CURRENT FILING DATE: 1995-06-05
EARLIER APPLICATION NUMBER: US 07/935,695
EARLIER FILING DATE: 1992-08-21
NUMBER OF SEQ ID NOS: 102
SOFTWARE: MS-word for Windows, Ver. 7.0
SEQ ID NO 8
LENGTH: 275
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
NAME/KEY: SCFY UHM8.1
LOCATION: 1..275
US-08-463-903-8

OTHER INFORMATION: :
US-07-935-695-8

Query Match 12.2%; Score 61; DB 4; Length 275;
Best Local Similarity 25.0%; Pred. No. 13;
Matches 22; Conservative 14; Mismatches 24; Indels 28; Gaps 4;

Db 8 AAGLLLLAQAOPAMADIVMSQSPSSLPVSGEKVTLSCSSQSLYSGNKNLYAWQOK 67

QY 9 ASAGALALLAAPAFQVY-----PVT--DELLANPAGEWISYCONQENY----- 51

Db 8 AAGLLLLAQAOPAMADIVMSQSPSSLPVSGEKVTLSCSSQSLYSGNKNLYAWQOK 67

QY 52 -RHSPITQITTEENVGOLQVWARGMOPG 78

Db 68 PGQSP-----KLLIYMASARESG 85

RESULT 7
US-08-463-903-10
Sequence 10, Application US/08463903
GENERAL INFORMATION:
APPLICANT: Mezes, Peter S.
APPLICANT: Richard, Ruth A.
APPLICANT: Affholter, Joseph A.
APPLICANT: Kollie, Nicolas J.
TITLE OF INVENTION: Dimer and Multimer Forms of Single Chain Polypeptides
FILE REFERENCE: 40224A US
CURRENT APPLICATION NUMBER: US/08/463,903
CURRENT FILING DATE: 1995-06-05
EARLIER APPLICATION NUMBER: US 07/935,695
EARLIER FILING DATE: 1992-08-21
NUMBER OF SEQ ID NOS: 102
SOFTWARE: MS-word for Windows, Ver. 7.0
SEQ ID NO 10
LENGTH: 280
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
NAME/KEY: SCFY UHM5.8
LOCATION: 1..280
US-08-463-903-10

Query Match 12.2%; Score 61; DB 3; Length 280;
Best Local Similarity 25.0%; Pred. No. 13;
Matches 22; Conservative 14; Mismatches 24; Indels 28; Gaps 4;

Db 8 AAGLLLLAQAOPAMADIVMSQSPSSLPVSGEKVTLSCSSQSLYSGNKNLYAWQOK 67

QY 9 ASAGALALLAAPAFQVY-----PVT--DELLANPAGEWISYCONQENY----- 51

Db 8 AAGLLLLAQAOPAMADIVMSQSPSSLPVSGEKVTLSCSSQSLYSGNKNLYAWQOK 67

QY 52 -RHSPITQITTEENVGOLQVWARGMOPG 78

Db 68 PGQSP-----KLLIYMASARESG 85

RESULT 8
US-07-935-695-10
Sequence 10, Application US/07935695
GENERAL INFORMATION:
APPLICANT: Mezes, Peter S.
APPLICANT: Richard, Ruth A.
APPLICANT: Affholter, Joseph A.
APPLICANT: Kollie, Nicolas J.
TITLE OF INVENTION: Dimer and Multimer Forms of Single Chain Polypeptides
FILE REFERENCE: 40224A US
CURRENT APPLICATION NUMBER: US/07/935,695
CURRENT FILING DATE: 1992-08-21
PRIOR APPLICATION NUMBER: US 08/463,903
PRIOR FILING DATE: 1995-06-05
NUMBER OF SEQ ID NOS: 102

```

; SOFTWARE: MS-Word for Windows, Ver. 7.0
; SEQ ID NO 10
; LENGTH: 280
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: SCFV UHM5.8
; LOCATION: 1..280
; OTHER INFORMATION:
US-07-935-695-10

```

```

Query Match
Best Local Similarity 12.2%; Score 61; DB 4; Length 280;
Matches 22; Conservative 14; Mismatches 24; Indels 28; Gaps 4;

```

```

OY 9 ASAGALLALAAPAFQVT-----PVT--DELLANPAGEMISYGNQENY----- 51
Db 8 AAAGLLLLAAQPMADIVMSQSPSSLPVSGEKVTLCKSSQSILYSGNKNYLAWYQOK 67
OY 52 -RHSPLTQITTEWVGQLVWARGMQPG 78
Db 68 PGQSP-----KLLIYWASARESG 85

```

```

RESULT 9
US-08-463-903-12
; Sequence 12, Application US/08463903
; Patent No. 6071515
; GENERAL INFORMATION:
; APPLICANT: Mezes, Peter S.
; APPLICANT: Richard, Ruth A.
; APPLICANT: Affholter, Joseph A.
; TITLE OF INVENTION: Dimer and Multimer Forms of Single Chain Polypeptides
; FILE REFERENCE: 40224A US
; CURRENT APPLICATION NUMBER: US/08/463,903
; EARLIER FILING DATE: 1995-06-05
; EARLIER FILING DATE: 1992-08-21
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: MS-Word for Windows, Ver. 7.0
; SEQ ID NO 12
; LENGTH: 282
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: SCFV UHM5.2
; LOCATION: 1..282
US-08-463-903-12

```

```

Query Match
Best Local Similarity 12.2%; Score 61; DB 3; Length 282;
Matches 22; Conservative 14; Mismatches 24; Indels 28; Gaps 4;

```

```

OY 9 ASAGALLALAAPAFQVT-----PVT--DELLANPAGEMISYGNQENY----- 51
Db 8 AAAGLLLLAAQPMADIVMSQSPSSLPVSGEKVTLCKSSQSILYSGNKNYLAWYQOK 67
OY 52 -RHSPLTQITTEWVGQLVWARGMQPG 78
Db 68 PGQSP-----KLLIYWASARESG 85

```

```

RESULT 10
US-07-935-695-12
; Sequence 12, Application US/07935695
; Patent No. 6329507
; GENERAL INFORMATION:
; APPLICANT: Mezes, Peter S.
; APPLICANT: Richard, Ruth A.
; APPLICANT: Affholter, Joseph A.

```

```

; APPLICANT: Kottle, Nicolas J.
; TITLE OF INVENTION: Dimer and Multimer Forms of Single Chain Polypeptides
; FILE REFERENCE: 40224A US
; CURRENT APPLICATION NUMBER: US/07/935,695
; CURRENT FILING DATE: 1992-08-21
; PRIOR APPLICATION NUMBER: US 08/463,903
; PRIOR FILING DATE: 1995-06-05
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: MS-Word for Windows, Ver. 7.0
; SEQ ID NO 12
; LENGTH: 282
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: SCFV UHM5.2
; LOCATION: 1..282
; OTHER INFORMATION:
US-07-935-695-12

```

```

Query Match
Best Local Similarity 12.2%; Score 61; DB 4; Length 282;
Matches 22; Conservative 14; Mismatches 24; Indels 28; Gaps 4;

```

```

OY 9 ASAGALLALAAPAFQVT-----PVT--DELLANPAGEMISYGNQENY----- 51
Db 8 AAAGLLLLAAQPMADIVMSQSPSSLPVSGEKVTLCKSSQSILYSGNKNYLAWYQOK 67
OY 52 -RHSPLTQITTEWVGQLVWARGMQPG 78
Db 68 PGQSP-----KLLIYWASARESG 85

```

```

RESULT 11
US-08-263-911-7
; Sequence 7, Application US/08263911
; Patent No. 5877291
; GENERAL INFORMATION:
; APPLICANT: Mezes, Peter S
; APPLICANT: Gourlie, Brian B
; TITLE OF INVENTION: MULTIVALENT SINGLE CHAIN ANTIBODIES
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Duane C. Ulmer
; STREET: P.O. Box 1967
; CITY: Midland
; STATE: MI
; COUNTRY: US
; ZIP: 48641-1967
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/263,911
; FILING DATE: 21-JUN-1994
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/990,263
; FILING DATE: 11-DEC-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Ulmer, Duane C
; REGISTRATION NUMBER: 34,941
; REFERENCE/DOCKET NUMBER: C-41,014
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (517) 636-8104
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 553 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear

```



```

; Patent No. 6066483
; GENERAL INFORMATION:
; APPLICANT: RIGGS, MICHAEL G.
; APPLICANT: SIVARAM, MATTHOK
; APPLICANT: TUDOR, STARLA D.
; TITLE OF INVENTION: PURIFIED DNA POLYMERASE FROM BACILLUS
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Gen-Probe Incorporated
; STREET: 9880 Campus Point Drive
; CITY: San Diego
; STATE: CA
; COUNTRY: USA
; ZIP: 92121
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/135,642
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/394,232
; FILING DATE:
; APPLICATION NUMBER: 08/307,410
; FILING DATE: 16-SEP-1994
; APPLICATION NUMBER: 08/222,612
; FILING DATE: 16-SEP-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Fisher, Carlos A
; REGISTRATION NUMBER: 36,510
; REFERENCE/DOCKET NUMBER: GP94003.CP2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-535-2807
; TELEFAX: 619-546-7929
; TELEX:
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 587 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEetical: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
; ORIGINAL SOURCE:
; US-09-135-642-23

Query Match          12.1%; Score 60.5; DB 3; Length 587;
Best Local Similarity 38.3%; Pred. No. 42;
Matches 18; Conservative 5; Mismatches 21; Indels 3; Gaps 1;

OY 17 LAAPAFQVTPYDELANPPAGEMISYGNQENYRHSPLTOTTEN 63
    || || |||||:|:| | | | | | | | | | | | | | |
Db 7 LAGMDFALADSVTDEMADKALVVEYV---DNYHAPIVGIALAN 50

RESULT 15
US-08-394-232a-23
; Sequence 23, Application US/08394232A
; Patent No. 6100078
; GENERAL INFORMATION:
; APPLICANT: RIGGS, MICHAEL G.
; APPLICANT: SIVARAM, MATTHOK
; APPLICANT: TUDOR, STARLA D.
; TITLE OF INVENTION: PURIFIED DNA POLYMERASE FROM BACILLUS STEAROTHEROPHILUS
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:

```

```

; ADDRESSEE: Gen-Probe Incorporated
; STREET: 9880 Campus Point Drive
; CITY: San Diego
; STATE: CA
; COUNTRY: USA
; ZIP: 92121
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/394,232A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/307,410
; FILING DATE: 16-SEP-1994
; APPLICATION NUMBER: 08/222,612
; FILING DATE: 16-SEP-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Fisher, Carlos A
; REGISTRATION NUMBER: 36,510
; REFERENCE/DOCKET NUMBER: GP94003.CP2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-535-2807
; TELEFAX: 619-546-7929
; TELEX:
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 587 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEtical: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
; ORIGINAL SOURCE:
; US-08-394-232a-23

Query Match          12.1%; Score 60.5; DB 3; Length 587;
Best Local Similarity 38.3%; Pred. No. 42;
Matches 18; Conservative 5; Mismatches 21; Indels 3; Gaps 1;

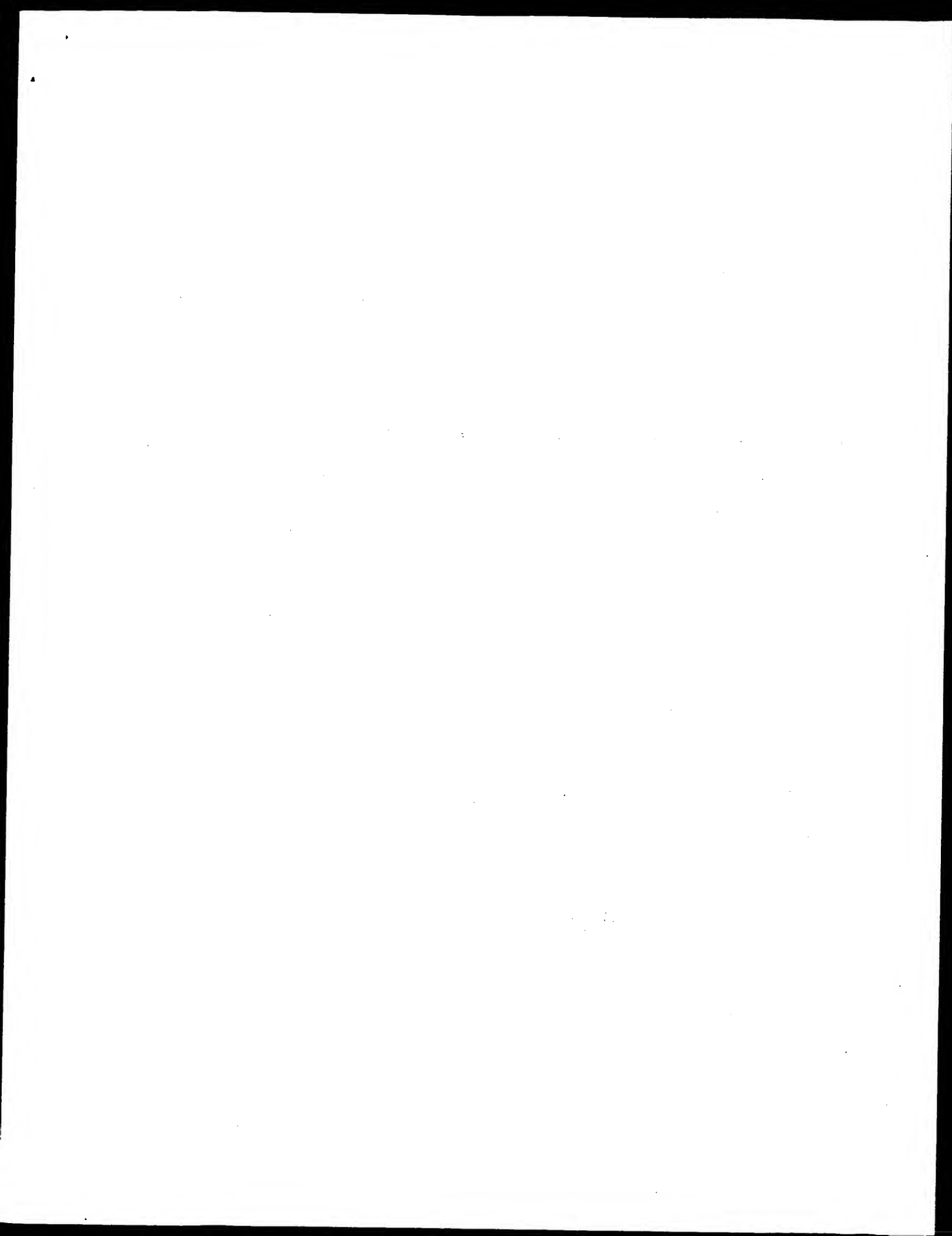
OY 17 LAAPAFQVTPYDELANPPAGEMISYGNQENYRHSPLTOTTEN 63
    || || |||||:|:| | | | | | | | | | | | | | |
Db 7 LAGMDFALADSVTDEMADKALVVEYV---DNYHAPIVGIALAN 50

```

Search completed: May 24, 2002, 10:02:54
 Job time: 228 sec

Fri May 24 11:27:21 2002

us-08-934-506a-5_copy_1_95.ra1



```

c/keywords: unbound
E:1-35/Domain: signal sequence #status predicted <SLIG>

```

F:36-738/Product: alcohol dehydrogenase 72K chain #status predicted <MAT>

Query Match 26.4%; Score 131.5; DB 2; Length 738;
Best Local Similarity 32.6%; Pred. No. 5.8e-06;

Matches 30; Conservative 15; Mismatches 38; Indels 9; Gaps 2;

OY 10 SAGALALLAAPAFQVPTVDE-----LLANPAGEWISYQONENYHSPFLQITTT 61
DB 17 TTTTCAALISGYATMASADGATGGAIIHADHPENMTYGRITTSYDQRYSPLDQINR 76
OY 62 ENVGQOLVWARGMQPGKQV-VPLIHGVMY 92
DB 77 SNGNKLKAWYLDLPTNRQESTPLVDGVMY 108

RESULT 3
AG0523

glucose dehydrogenase [imported] - Salmonella enterica subsp. enterica serovar Typh1 (st
C:Species: Salmonella enterica subsp. enterica serovar Typh1

A:Note: this species has also been called Salmonella typhi

C:Date: 09-Nov-2001 #sequence,revision 09-Nov-2001 #text_change 27-Nov-2001

C:Accession: AG0523

R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
th, T.; Connerly, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
S.; Moule, S.; O'Gaora, P.

Nature 413, 848-852, 2001

A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;

A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov

A:Reference number: AB0502; PMID:11677608

A:Accession: AG0523

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-796 <PAR>

A:Cross-references: GB:AL513382; PIDN:CAD01327.1; PID:g16501455; GSPDB:GN00176

C:Genetics:

A:Gene: STR0191

C:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)

Query Match 26.0%; Score 129.5; DB 2; Length 796;
Best Local Similarity 34.6%; Pred. No. 6.8e-06;

Matches 36; Conservative 14; Mismatches 29; Indels 25; Gaps 5;

OY 6 LWAS-----AGALALLAAPAFQVPTVDELLANPAGEWISYQONENYHSPFLQ 58
DB 135 LTMAGFNDPQINGTLRADATPA-ATSSSIAD-----DPAVGRNGEGGRYSPLKO 185
OY 59 ITTENVGOLVWV-----RCMQPGKV--CVPLIHGVMY 93
DB 186 ITADNVHQLKEAVFRTGLKQPNDEITNEVTPIKGDTLYL 229

RESULT 4
J50326

alcohol dehydrogenase (EC 1.1.1.7) 72K chain precursor - Acetobacter aceti
C:Species: Acetobacter aceti

C:Date: 31-Mar-1990 #sequence,revision 31-Mar-1990 #text_change 20-Jun-2000

C:Accession: J50326

R:Inoue, T.; Sunagawa, M.; Mori, A.; Imai, C.; Fukuda, M.; Takagi, M.; Yano, K.

J. Bacteriol. 171, 3115-3122, 1989

A:Title: Cloning and sequencing of the gene encoding the 72-kilodalton dehydrogenase sub

A:Reference number: J50326; MUID:89255070

A:Accession: J50326

A:Molecule type: DNA

A:Residues: 1-742 <INO>

A:Cross-references: GB:D90004; GB:M26951; NID:g216193; PIDN:BAAL4058.1; PID:g216194

A:Experimental source: Strain K6033

A:Note: amino terminal of mature protein is confirmed

C:Genetics:

A:Gene: adh1

C:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)

C:Keywords: alcohol metabolism; NAD; oxidoreductase

F:135/Domain: signal sequence #status predicted <SIG>

F:36-742/Product: alcohol dehydrogenase 72K dehydrogenase chain #status predicted <AD

Query Match 25.6%; Score 127.5; DB 2; Length 742;
Best Local Similarity 38.8%; Pred. No. 1e-05;

Matches 33; Conservative 11; Mismatches 38; Indels 3; Gaps 2;

OY 11 AGALALLAAPAFQVPTVDELL--ANPAGEWISYQONENYHSPFLQITTENYGO 68
DB 23 AAALPYAANPARADQGTGGAIIHADHPENMTYGRITTSYDQRYSPLDQINR 82
OY 69 LVMARGMQPGKQV-VPLIHGVMY 92
DB 83 LGGYTLPTNRQESTPLVDGVMY 107

RESULT 5
JY0107

glucose dehydrogenase (pyrroloquinoline-quinone) (EC 1.1.99.17) - Escherichia coli

C:Species: Escherichia coli

C:Date: 30-Jun-1992 #sequence,revision 12-Dec-1997 #text_change 16-Jun-2000

C:Accession: D64735; JY0107; A45997; S45201; I41228

R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.;

A.; Rose, D.J.; Mau, B.; Shao, Y.

Science 277, 1453-1462, 1997

A:Title: The complete genome sequence of Escherichia coli K-12.

A:Reference number: A64720; MUID:97426617

A:Accession: D64735

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-796 <BLAT>

A:Cross-references: GB:AE000122; GB:U00096; NID:g1786315; PIDN:AAAC73235.1; PID:g17863

A:Experimental source: strain K-12, substrain M61655

R:Cleton-Jansen, A.M.; Goosen, M.; Fayet, O.; van de Putte, P.

J. Bacteriol. 172, 6308-6315, 1990

A:Title: Cloning, mapping, and sequencing of the gene encoding Escherichia coli quino

A:Reference number: JY0107; MUID:91035240

A:Accession: JY0107

A:Molecule type: DNA

A:Residues: 1-58, 'L', 60-148, 'H', 150, 'K', 194-796 <CLE>

A:Cross-references: GB:X51323; NID:g41553; PIDN:CAAS5706.1; PID:g41554

A:Experimental source: strain K12

R:Yamada, M.; Sumi, K.; Matsushita, K.; Adachi, O.; Yamada, Y.

J. Biol. Chem. 268, 12812-12817, 1993

A:Title: Topological analysis of quinoprotein glucose dehydrogenase in Escherichia co

A:Reference number: A45997; MUID:93286127

A:Accession: A45997

A:Status: preliminary

A:Molecule type: protein

A:Residues: 1-20 <YAM>

R:Fujita, N.

submitted to the EMBL Data Library, January 1994

A:Reference number: S45181

A:Accession: S45201

A:Molecule type: DNA

A:Residues: 1-796 <FUJ>

A:Cross-references: EMBL:D26562; NID:g473770; PIDN:BA05580.1; PID:g473791

A:Experimental source: strain K-12 substrain W3110

R:Yamada, M.; Asaoka, S.; Saiter, M.H.; Yamada, Y.

J. Bacteriol. 175, 568-571, 1993

A:Title: Characterization of the gcd gene from Escherichia coli K-12 W3110 and regula

A:Reference number: I41228; MUID:93123180

A:Accession: I41228

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-148, 'H', 150, 'K', 194-665, 'H', 667-796 <RES>

A:Cross-references: GB:DL2651; NID:g216555; PIDN:BA02174.1; PID:g216556

C:Genetics:

A:Gene: gcd

A:Map position: 3 min

C:Function:

A:Description: catalyzes the oxidation of D-glucose to D-gluconic acid by ubiquinone

A:Pathway: respiratory chain
C:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)
Keywords: oxido-reductase; pyrroloquinoline quinone; respiratory chain; transmembrane F
F:11-37/Domain: transmembrane #status predicted <TM1>
F:41-59/Domain: transmembrane #status predicted <TM2>
F:63-81/Domain: transmembrane #status predicted <TM3>
F:96-110/Domain: transmembrane #status predicted <TM4>
F:120-140/Domain: transmembrane #status predicted <TM5>
F:193-95/Binding site: ubiquinone (Arg, Asp) #status predicted
F:466/Active site: Asp #status predicted

Query Match	25.6%;	Score 127.5;	DB 1,	Length 796;
Best Local Similarity	33.7%;	Pred. No. 1.le-05;		
Matches	35;	Conservative	15;	Mismatches 29;
			Indels	25;
			Gaps	5

QY 6 LTMAS-----AGALLIAPAFADQVTPVTDELLNPPAGMWTISYGONQENYHPSLTQ 58
 | | : : | | : : | : : : : : : : : : : : : : : : : :
Db 135 LTWAGNDPQEIINGTLSDATPAEA-ISPVDQ-----DMPAYGNQGSGQFSPKQ 185

```
QY      59 ITTENVGQLOLWMA-----KGMOPKAV-QVTPPLHIDGVAKL  93  
       |::||:|           : ||::|||:   :||  
Db     186 INADVNHLKEAWERTGDVKQPNDPGETITNEVTPIKVGDTLYL  229
```

RESULT 6
H90644
glucose dehydrogenase [imported] - Escherichia coli (strain O157:H7, substrain RIMD 0503
C:Species: Escherichia coli
#text change 03-Aug-2001
revision 18-Jul-2001

C;Accession: H90644
R;Hayashi, T.; Makino, K.; Onishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and genome
Reference number: A99629; M01D:21156231; PMID:11258796

A:Accession: J90644
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-796 <NAV>
A:Cross-references: GB:AA000007; PID: BAB3551.1; PID: I41359584; GSPDB: G0000154

A: Experimental source: strain 0157:H7, substrain N1MD 050352
C: Genetics:
A: Gene: ECs0128
C: Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)

Query Match 25.6%; Score 127.5; DB 2; Length 796;
Best Local Similarity 33.7%; Pred. No. 1.1e-05;
Matches 35; Conservative 15; Mismatches 29; Indels 25; Gaps 5.

Oy 6 LWMAS-----AGALLALPAFAQVPTVTELLANPAGEMISYQONERYHSPLTQ 58
 | | | | | : | | | : | : | | | | | | | |
 Dn 135 LTWAGNDPOEINGTSLADATPHEA-ITSPVADQ-----DMPATGRNQEQRFSP LKQ 185

59 ITTENVGOLYLWVA-----RGMOPGV- QYPTPLIHGVMYL 93
 1 111 1 : 111 : 111 : 111
 Db 186 IINDNVINKEMVFRFTGDYKQPNPGEITNEVTPPIKVGDTLYL 229

```

RESULT      7
H85495
glucose dehydrogenase [imported] - Escherichia coli (strain O157:H7, substrain EDL933)
C:Species: Escherichia coli
C:Date: 16-Feb-2001 #sequence revision 16-Feb-2001 #text-change 14-Sep-2001

```

C. Bacter: 10.1007/s00284-001-0001-0
C; Accession: H85495
N; Bacteria, N.T.; Plunnebeck III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
N; Bacteria, N.T.; Plunnebeck III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grobbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A. r11a: Genome sequence of enterohemorrhagic *Escherichia coli* O157:H7.

A;Reference number: A85480; MID:21074935; PMID:11206551
A;Accession: H85495
A;Status: preliminary

A:Molecule type: DNA
A:Residues: 1-796 <STO>
A:Cross-references: GB:AE005174, NID:912512839, PIDN:AA654428.1; GSPDB:GN00145; UMGF
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: gcd
C:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)

Query Match	25.6%;	Score 127.5;	DB 2;	Length 796;
Best Local Similarity	33.7%;	Pred. No. 1.1e-05;		
Matches	35;	Conservative	15;	Mismatches 29;
				Indels 25;
				Gaps 5;

QY 6 LIMS-----ACALALLAPAPAQTPTVDLLANPAGENISYQONCEHTSHSLTQ 58
 | | | : | | | : | | : : | : | | | | | | |
Db 135 LTWAGFNDPQETINGTISADATPAEA-ISPVAQ-----DMPAYIGNGEGGKFSPLKQ 185

```

QY      59 ITTENVGQLOLWMA-----RGNQPGKY-QYTPPLHIDGVMTL 93
      | :| | : |      : | | | : | |
Db     186 INADNHNLEAWEFRTGDYKQPNPDGSEITNEVPIKVGDTLYL 229

```

RESULTS
S53317
guinohemoprotein ethanal dehydrogenase (EC 1.1.99.-) type 1 precursor - Comamonas test
G:cherries: Comamonas testosteroni

C:\Species\1995\#sequence_revision 21-Jul-1995 #text_change 02-Jun-2000
C:\Date: 08-May-1995 #sequence_revision 21-Jul-1995 #text_change 02-Jun-2000
C:\Accession: S62366; S62373; S65908; S52317
R:\Biochem\J.J. Kraayveld, D.E.; van Sluys, C.A.; Jongejan, J.A.; de Vries, S.; Duijn
R:\Biochem\J.J. Kraayveld, D.E.; van Sluys, C.A.; Jongejan, J.A.; de Vries, S.; Duijn
Eur. J. Biochem. 235, 690-698, 1996
A:\Title: Characterization of the gene encoding quinohemoprotein ethanol dehydrogenase
A:\Reference number: S62366; M010: 96184549

A:Accession: S62366
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-708 <ST01>
A:Cross-references: EMBL:X81880; NID:g663195; PIDD:CAA57464.1; PID:g663196

A:Accession: S62373
A:Status: preliminary
A:Molecule type: protein
A:Residues: 32-54 <TO>
R:de Jong, G.A.H.; Geerlof, A.; Stoorvogel, J.; Jondejan, J.A.; de Vries, S.; Duine,

Eur. J. Biochem. 230, 899-905, 1995
A>Title: Quinoloneprotein ethanol dehydrogenase from *Comamonas testosteroni*. Purified
A;Reference number: S65908; MUID:95324580
A;Accession: S65908

A: Molecule type: protein
A: Residues: 32-50, 'X', 52-54: 477-483, 'X', 485-490 <DEU>
A: Experimental source: ATCC 15667
C: Genetics:
A: Gene: qheh

C:Superfamily glucose dehydrogenase (pyridoxaldehyde dependent)
C:Superfamily glucose dehydrogenase
C:Keywords: oxidoreductase; quinoprotein
E:1-31/Domain: signal sequence #status predicted <Sig>
F:1-3708/product: quinoximoprotein ethanol dehydrogenase type I #status experimental

Query Match	23.8%;	Score 118.5;	DB 2;	Length 708;
Best Local Similarity	30.5%;	Pred. No. 9.2e-05;		
	10.0;	Mismatches 40;		
		Indels 15;		
		Gaps 4;		

Qy 3 PTSLTMSAGALALLMAPAQOVTPT -----DELTANP-AGEVSTYQONEN 50
::: ||| |
Db 12 PGRWTLIACLG--SAAAPAGTGPPAQAAYAVRDGCFIRNNAARTPDWTFIVGYDAE 69

```

OY      31  IRKSPFLVILINVCQWZLNHSSGSGS : : : : : : : : : : : : : : : :
Db       70  TRYSRLDGINAANVKDLGLAMSTYNDESTRGVEATPYVDVGIMYVS 114

```

glucose dehydrogenase PA230 [imported] - Agrobacterium tumefaciens (strain C58, Cere

	Query Match	22.4%	Score 111.5;	DB 2;	Length 778;	
	Best Local Similarity	28.7%;	Pred. No. 0.00059;			
	Matches	35;	Conservative	15;	Mismatches 39;	Indels 33; Gaps 5
QY	3	PTSLWA-SAGALALAAAPA-----	-GVPTDELLANPPA-----	GERT	42	
		: : : :	: : : :	:	:	
Db	108	PTGLSPNAMPVLVLSLFAVAAGYSMAODPHDAGSLPELTASAPVYGVEPDGWH			167	
QY	43	SYSONCENTRHSEPLQTITTENVGOQLVMWAKMGPGKV-----	-GVPLINDGYV		91	
		: : : :		:	:	
Db	168	QYGRTEYCGRYSLPQOVNVNDVNSQLKEAMR--YTGDVKLPDDVDGETTYGVTLKLGNTLI			225	

RESULT 12
OPREX
glucose dehydrogenase (pyrroloquinoline-quinone) (EC 1.1.99.17) - Gluconobacter oxydans
C:Species: Gluconobacter oxydans
C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 22-May-1998
C:Accession: S17716; S19265
R:Cleton-Jansen, A.M.; Dekker, S.; van de Putte, P.; Goosen, N.
Mol. Gen. Genet. 229, 206-212, 1991
A:Title: A single amino acid substitution changes the substrate specificity of quinop
A:Reference number: S17716; MIMD:92017653
A:Accession: S17716
A:Molecule type: DNA
A:Residues: 1-808 <CLE>
A:Cross-references: EMBL:X62710
R:Goosen, N.
Submitted to the EMBL Data Library, February 1992
A:Reference number: S19265
A:Accession: S19265
A:Molecule type: DNA
A:Residues: 1-212, 'A', 214-808 <GOO>
A:Cross-references: EMBL:X62710; NID:g58416; PID:g58417
C:Genetics:
A:Gene: gdh
C:Function:
A:Description: catalyzes the oxidation of D-glucose to D-gluconic acid by ubiquinone
A:Pathway: respiratory chain
C:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)

C:Keywords: oxidoreductase; pyrroloquinoline quinone; respiratory chain; transmembrane
 F:9-28/Domain: transmembrane #status predicted <TM1>
 F:35-54/Domain: transmembrane #status predicted <TM2>
 F:60-76/Domain: transmembrane #status predicted <TM3>
 F:94-110/Domain: transmembrane #status predicted <TM4>
 F:122-138/Domain: transmembrane #status predicted <TM5>
 F:91,93/Binding site: ubiquinone (Ar9, Asp) #status predicted
 F:470/Active site: Asp #status predicted

Query Match 21.7%; Score 108; DB 1; Length 808;
 Best Local Similarity 28.2%; Pred. No. 0.0015;
 Matches 29; Conservative 14; Mismatches 40; Indels 20; Gaps 3;

13 ALALLAFAFAVPTVDEL-----NPPAGEWISYGNQENTRHSPLOTITT 61
 128 ALVALLAFSLFTPDHDSGLPTQIANASPADDPNVPASPMHAYGRQAGDRWSPINQINA 187
 62 ENVGQLQVW-----ARGMQPGKV--QVTPLIHDGVWYLAN 95
 188 TTVSNLKVAMHHTKDMNNSNDPEQTNEATPIEFNNITLMCS 230

RESULT 13

E95863
 Alcohol dehydrogenase (acceptor) (EC 1.1.99.8) large subunit [Imported] - Sinorhizobium

C:Species: Sinorhizobium meliloti
 C:Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 14-Sep-2001

C:Accession: E95863
 R:Finan, T.M.; Weidner, S.; Wong, K.; Buhmester, J.; Chain, P.; Vorholter, F.J.; Hernar

Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001
 A:Title: The complete sequence of the 1,683-kb pSymB megaplasmid from the N2-fixing endo

A:Reference number: A95842; MUID:21396508; PMID:11481431

A:Accession: E95863

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-601 <KUR>

A:Cross-references: GB:AL591985; PIDN:CAC48573.1; PID:g15140045; GSPDB:GN00167

A:Experimental source: strain 1021, megaplasmid pSymB

R:Galibert, F.; Finan, T.M.; Long, S.R.; Publer, A.; Abola, P.; Ampe, F.; Barloy-Hubler, J.;

pell, D.; Chain, P.; Cowle, A.; Davis, R.W.; Dreano, S.; Federpiet, N.A.; Fisher, R.F.;

L.; Hyman, R.W.; Jones, T.
 Science 293, 668-672, 2001

A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,

hehault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.

A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.

A:Reference number: A96039; MUID:21368234; PMID:11474104

A:Contents: annotation

C:Genetics:

A:Gene: SMD20173

A:Genome: Plasmid

C:Keywords: oxidoreductase

Query Match 19.9%; Score 99; DB 2; Length 601;
 Best Local Similarity 34.5%; Pred. No. 0.0097;
 Matches 29; Conservative 17; Mismatches 32; Indels 6; Gaps 4;

14 LALLAFAFAVPTVDEL--LANPAGEWISYGNQENTRHSPLOTITTENVGOLQVW 71
 8 LAIWSIGGAOYAFANDELQKLIIDP--NOMAIQIGDYANLRYSKLDQINKNVCKLOVAV 66
 72 --ARGMQPGKVQVTPLIHDGVWYL 93
 67 TSTGVLRGH--EGSPLYIGDLAVY 89

RESULT 14

S00943
 glucose dehydrogenase (pyrroloquinoline-quinone) (EC 1.1.99.17) - Acinetobacter calcoac

R:Cleton-Jansen, A.M.; Goosen, N.; Odle, G.; van de Putte, P.

Nucleic Acids Res. 16, 6228, 1988

A:Title: Nucleotide sequence of the gene coding for quinoprotein glucose dehydrogenas

A:Reference number: S00943; MUID:88289368

A:Accession: S00943

A:Molecule type: DNA

A:Residues: 1-801 <CLE>

A:Cross-references: EMBL:X07235; NID:g38711; PIDN:CAA30222.1; PID:g38712

A:Experimental source: strain LMD 79.41

C:Function:

A:Description: catalyzes the oxidation of D-glucose to D-gluconic acid by ubiquinone

A:Pathway: respiratory chain

C:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)

C:Keywords: oxidoreductase; pyrroloquinoline quinone; respiratory chain; transmembran

F:9-35/Domain: transmembrane #status predicted <TM1>
 F:39-57/Domain: transmembrane #status predicted <TM2>
 F:61-79/Domain: transmembrane #status predicted <TM3>
 F:94-108/Domain: transmembrane #status predicted <TM4>
 F:118-137/Domain: transmembrane #status predicted <TM5>
 F:91,93/Binding site: ubiquinone (Ar9, Asp) #status predicted
 F:471/Active site: Asp #status predicted

Query Match 19.1%; Score 95; DB 1; Length 801;
 Best Local Similarity 31.2%; Pred. No. 0.0372;
 Matches 24; Conservative 11; Mismatches 22; Indels 20; Gaps 3;

20 PAFACVPTVDELALNPPAGEWISYGNQENTRHSPLOTITTENVGOLQVWARGMQPGK 79
 152 PETAQAVGVAB-----SDMPAYGRQAGVRSPLQINDQNVKDLKAVMT--LRTGD 202
 80 V-----QVPL 85
 203 LKTDNDGSETTQVPTI 219

RESULT 15

J00706
 Alcohol dehydrogenase (acceptor) (EC 1.1.99.8) alpha chain precursor - Methylobacteri

M:Alternate names: methanol dehydrogenase 62K large chain

C:Species: Methylobacterium extorquens

C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 08-Oct-1999

C:Accession: J00706

R:Anderson, D.J.; Morris, C.J.; Nunn, D.N.; Anthony, C.; Lidstrom, M.E.

Gene 90, 173-176, 1990

A:Title: Nucleotide sequence of the Methylobacterium extorquens AM1 mox and mox2 gen

A:Reference number: J00706; MUID:90337342

A:Accession: J00706

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-626 <AND>

A:Cross-references: GB:M31108; NID:g150017; PIDN:AAA25380.1; PID:g150018

A:Experimental source: strain AM1

Biochem. J. 260, 857-862, 1989

A:Title: The second subunit of methanol dehydrogenase of Methylobacterium extorquens

A:Reference number: S04644; MUID:89350892

A:Accession: S07908

A:Molecule type: protein

A:Residues: 28-50, 'XX', 53 <NN>

A:Note: the source is designated as Methylobacterium extorquens AM1

C:Comment: This enzyme oxidizes methanol to formaldehyde.

C:Genetics:

A:Gene: moxP

C:Keywords: alcohol metabolism; oxidoreductase

F:1-27/Domain: signal sequence #status predicted <Sig>

F:28-626/Product: alcohol dehydrogenase (acceptor) #status experimental <Mat>

Query Match 18.9%; Score 94; DB 2; Length 626;
 Best Local Similarity 34.8%; Pred. No. 0.036;

Matches 31; Conservative 16; Mismatches 36; Indels 6; Gaps 5;

QY 10 SAGALLLA-AP-AFAQVTPVTDELIANPPAGE-WISYGONENYRHSPLTOITTEENYGO 66
 Db 7 SVSALAMIALAPALSSGAYANDKLELSKSDDNWVMPGKNYDSNNFSDPKOINKGNVYO 66
 QY 67 LQLYW--ARGMOPGKYQVTPPLIHGVMYL 93
 Db 67 LRPAMTFSTGLNGH-EGAPLYVDGKMYI 94

Search completed: May 24, 2002, 10:06:02
 Job time: 286 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2002, 10:03:02 ; Search time 27.86 Seconds
(without alignments)
132.030 Million cell updates/sec

Title: US-08-934-506a-5_COPY_1_95
Perfect score: 498
Sequence: 1 MKPTSLWASAGALALAAP.....QPKVQVTPLIHQVATLAN 95

Scoring table: BIOSIM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 segs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	158.5	31.8	757	1	DHET_GLUOX
2	131.5	26.4	738	1	DHET_ACEPO
3	130.5	26.2	739	1	DHET_ACEBU
4	127.5	25.6	742	1	DHET_ACEAC
5	127.5	25.6	796	1	DHG_ECOLI
6	108	21.7	808	1	DHG_GLUOX
7	104	20.9	626	1	DHML_MENOR
8	95	19.1	801	1	DHGA_LACTA
9	94	18.9	626	1	DHML_METEX
10	93	18.3	631	1	DHML_PAREE
11	91	18.3	623	1	EXXA_PSEAE
12	76	15.3	733	1	TNFA_MARMO
13	76	15.3	733	1	PSAB_ODOSI
14	72.5	14.6	776	1	XOAF_HETTR
15	70.5	14.2	600	1	HEWZ_SEUMA
16	68.5	13.8	417	1	HEWZ_THERU
17	68.5	13.8	441	1	PSAB_PROHO
18	65.5	13.2	661	1	GSPN_PSEAE
19	64	12.9	235	1	GCL_RAT
20	64	12.9	326	1	GCL_RAT
21	63.5	12.8	369	1	YDHH_ECOLI
22	63	12.7	885	1	P77570 escherichia
23	62.5	12.6	128	1	YD03_YEAST
24	62.5	12.6	134	1	RS11_BACST
25	62	12.4	310	1	RS11_BORBU
26	62	12.4	322	1	MM04_ECOLI
27	62	12.4	453	1	GCA_RAT
28	62	12.4	4367	1	Y259_MYCPN
29	61.5	12.3	148	1	DYHC_NEUCR
30	61.5	12.3	221	1	CY08_LOCOMI
31	61	12.2	247	1	XYNI_COCCA
32	61	12.2	412	1	YCIC_ECOLI
33	61	12.2	735	1	HEW2_SOYBN
					PSAB_CHLMO

34	61	12.2	753	1	METE_SALTY	O916n1 salmonella
35	61	12.2	767	1	HYFE_SYNY3	O55638 synchocyst
36	61	12.2	864	1	STAZ_PIG	O02799 sus scrofa
37	60.5	12.1	876	1	DP01_BACST	P20206 bacillus st
38	59.5	11.9	139	1	RS11_BACST	P04969 bacillus su
39	59.5	11.9	139	1	RS11_MYCHO	P45812 mycobacteri
40	59.5	11.9	146	1	AZ0P_MLCRA	P43377 alcaigenes
41	59.5	11.9	221	1	SEGA_BPT4	P32286 bacterioph
42	59	11.8	176	1	RNMG_ASPEE	P04389 aspergillus
43	59	11.8	233	1	RLI_SERMA	P09764 serratia ma
44	59	11.8	270	1	CB12_LYCES	P10708 lycopersico
45	59	11.8	300	1	MOXI_METEX	P16028 methylolact

ALIGNMENTS

RESULT	ID	DEHT_GLUOX	STANDARD	PRT	757 AA
AC	005542				
DE	01-NOV-1997 (Rel. 35, Created)				
DE	01-NOV-1997 (Rel. 35, Last sequence update)				
DE	16-OCT-2001 (Rel. 40, Last annotation update)				
DE	Alcohol dehydrogenase [acceptor] precursor (EC 1.1.99.8) (G3-ADH subunit 1).				
GN	ADHA.				
OS	Glucobacter oxydans (Glucobacter suboxydans).				
OC	Bacteria; Proteobacteria; alpha subdivision; Acetobacteraceae; Glucobacter.				
OX	NCBI_TaxID=442;				
RM	[1]				
RP	SEQUENCE FROM N.A., AND SEQUENCE OF 36-50.				
RC	STRAIN-IFO 12528;				
RX	MEDLINE=97208225; PubMed=9055427;				
RA	Kondo K., Horinouchi S.;				
RT	"Characterization of the genes encoding the three-component membrane-bound alcohol dehydrogenase from Glucobacter suboxydans and their expression in Acetobacter pasteurianus."				
RT	Appl. Environ. Microbiol. 63:1131-1138(1997).				
RL	FUNCTION: DEHYDRATION OF PRIMARY ALCOHOLS (EXCEPTION: METHANOL) (BY SIMILARITY).				
CC	-1- CATALYTIC ACTIVITY: A primary alcohol + acceptor = an aldehyde + reduced acceptor.				
CC	-1- COFACTOR: PO4 AND HEME.				
CC	-1- SUBUNIT: HETERODIMER (DEHYDROGENASE, CYTOCHROME AND PROTEIN ADHS), THAT FORMS THE ALCOHOL DEHYDROGENASE COMPLEX.				
CC	-1- SUBCELLULAR LOCATION: MEMBRANE-BOUND, FACING THE PERIPLASMIC SPACE (POTENTIAL).				
CC	-1- SIMILARITY: BELONGS TO THE BACTERIAL PO4 DEHYDROGENASE FAMILY.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).				
CC	EMBL: D86375; BAA19753.1; -				
DR	HSSP: Q97437; 1F1G.				
DR	InterPro: IPR001479; Bac_P00.				
DR	InterPro: IPR002372; Bac_P00_repeat.				
DR	InterPro: IPR000345; CytC_heme_bind.				
DR	Pfam: PF01011; Bacterial_P00_1; 1.				
DR	PROSITE: PS00363; BACTERIAL_P00_1; 1.				
DR	PROSITE: PS00364; BACTERIAL_P00_2; 1.				
DR	PROSITE: PS00190; CYTOCHROME_C; 1.				
DR	Oxidoreductase; PO4; Heme; Periplasmic; Membrane; Signal.				
KM	SIGNAL	34			
FT	CHAIN	35	757		
FT	MOD_RES	35	35		
FT	ACT_SITE	342	342		
FT					ALCOHOL DEHYDROGENASE [ACCEPTOR].
FT					PYRROLIDONE CARBOXYLIC ACID.
FT					BASE (POTENTIAL).

FT BINDING 653 653 HEME (COVALENT) (BY SIMILARITY).
 FT BINDING 656 656 HEME (COVALENT) (BY SIMILARITY).
 FT METAL 657 657 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
 SQ SEQUENCE 757 AA: 82968 MW: 3989F90E3B947581 CRC64.

Query Match 31.8%; Score 158.5; DB 1; Length 757;
 Best Local Similarity 45.7%; Pred. No. 3.5e-09;
 Matches 42; Conservative 12; Mismatches 31; Indels 7; Gaps 4;

OY 7 LMSAGALALLAA--PFAOV--TPVTDELANPPAGEWISYGOONENRHSPLTQIT 61
 DB 16 LLSCAALAFSAVPAFAQEDTGTAITSNDGHP-CDWLSYSGNSYEQXYSPLDQINT 74
 OY 62 ENVGOLQLYMANGMOPGKQV--TPLIHDCVMY 92
 DB 75 ENVGKLKLMHYDLDTNRGQESTPLIVGVMT 106

RESULT 2

DHET_ACEPO
 ID DHET_ACEPO STANDARD: PRT: 738 AA.

AC P28036:
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Alcohol dehydrogenase [acceptor] precursor (EC 1.1.99.8).
 GN ADHA.
 OS Acetobacter polyoxogenes.
 OC Bacteria; Proteobacteria; alpha subdivision; Acetobacteraceae;
 CC Acetobacter.
 OX NCBI_TaxID=439;
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC STRAIN=NB11028;
 RX MEDLINE=91159482; Pubmed=2001402;
 RA Tamaki T., Fukaya M., Takemura H., Tayama K., Okumura H.,
 RA Kawamura Y., Nishiyama M., Horiuchi S., Beppu T.;
 RT "Cloning and sequencing of the gene cluster encoding two subunits of
 RT membrane-bound alcohol dehydrogenase from Acetobacter polyoxogenes.,"
 RL Biochim. Biophys. Acta 1088:292-300(1991).
 CC -1- FUNCTION: DEHYDRATION OF PRIMARY ALCOHOLS (EXCEPT: METHANOL).
 CC -1- CATALYTIC ACTIVITY: A primary alcohol + acceptor = an aldehyde +
 CC reduced acceptor.
 CC -1- COFACTOR: PQO AND HEME.
 CC -1- SUBUNIT: HETEROTETRAMER (DEHYDROGENASE, CYTOCHROME, AND TWO
 CC SMALLER UNKNOWN SUBUNITS) THAT FORMS THE ALCOHOL DEHYDROGENASE
 CC COMPLEX.
 CC -1- SUBCELLULAR LOCATION: MEMBRANE-BOUND, FACING THE PERIPLASMIC
 CC SPACE (POTENTIAL).
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL PQO DEHYDROGENASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: D00635; BAA00528.1; -;
 DR PIR: S14270; S14270.
 DR HSSP: G92477; 1TFG.
 DR InterPro: IPR001479; Bac_PQO.
 DR InterPro: IPR002372; Bac_PQO_repeat.
 DR Pfam: PF01011; Bacterial_PQO_6.
 DR PROSITE: PS00363; BACTERIAL_PQO_1; 1.
 DR PROSITE: PS00364; BACTERIAL_PQO_2; 1.
 DR PROSITE: PS00190; CYTOCHROME_C; 1.
 KW Oxidoreductase; PQO; Heme; Periplasmic; Membrane; Signal.
 FT SIGNAL 1 35
 FT CHAIN 36 738
 FT POTENTIAL.
 FT ALCOHOL DEHYDROGENASE [ACCEPTOR].

FT BINDING 650 650 HEME (COVALENT) (BY SIMILARITY).
 FT BINDING 653 653 HEME (COVALENT) (BY SIMILARITY).
 FT METAL 654 654 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
 SQ SEQUENCE 738 AA: 80840 MW: 1E2B6ED7BCD92AF6 CRC64.

Query Match 26.4%; Score 131.5; DB 1; Length 738;
 Best Local Similarity 32.6%; Pred. No. 2.5e-06;
 Matches 30; Conservative 15; Mismatches 38; Indels 9; Gaps 2;

OY 10 SAGALALLAAPFAQVTPVTDE-----LLANPAGEWISYGOONENRHSPLTQIT 61
 DB 17 TAGTICALLISGVATMASADGAGTGAIIHADHPENMMYTRITTSORTSPLDQINR 76
 OY 62 ENVGOLQLYMANGMOPGKQV--TPLIHDCVMY 92
 DB 77 SNVGKLKLMHYDLDTNRGQESTPLIVGVMT 108

RESULT 3

DHET_ACEEU
 ID DHET_ACEEU STANDARD: PRT: 739 AA.

AC Q44002:007952;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Alcohol dehydrogenase [acceptor] precursor (EC 1.1.99.8).
 GN ADH.
 OS Acetobacter europaeus.
 OC Bacteria; Proteobacteria; alpha subdivision; Acetobacteraceae;
 CC Gluconacetobacter.
 OX NCBI_TaxID=33995;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=DES11 / DSM 6160;
 RX Thurner C.A.K.;
 RA Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
 RL -1- CATALYTIC ACTIVITY: A primary alcohol + acceptor = an aldehyde +
 CC reduced acceptor.
 CC -1- COFACTOR: PQO AND HEME (BY SIMILARITY).
 CC -1- SUBUNIT: HETEROTETRAMER (DEHYDROGENASE, CYTOCHROME, AND TWO
 CC SMALLER UNKNOWN SUBUNITS) THAT FORMS THE ALCOHOL DEHYDROGENASE
 CC COMPLEX (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: MEMBRANE-BOUND, FACING THE PERIPLASMIC
 CC SPACE (POTENTIAL).
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL PQO DEHYDROGENASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: X82894; CAA58066.1; -;
 DR EMBL: Y09480; CAA70688.1; -;
 DR HSSP: G92477; 1TFG.
 DR InterPro: IPR001479; Bac_PQO.
 DR InterPro: IPR002372; Bac_PQO_repeat.
 DR Pfam: PF01011; Bacterial_PQO_6.
 DR PROSITE: PS00363; BACTERIAL_PQO_1; 1.
 DR PROSITE: PS00364; BACTERIAL_PQO_2; 1.
 DR PROSITE: PS00190; CYTOCHROME_C; 1.
 KW Oxidoreductase; PQO; Heme; Periplasmic; Membrane; Signal.
 FT SIGNAL 1 35
 FT CHAIN 36 739
 FT BINDING 651 651 HEME (COVALENT) (BY SIMILARITY).
 FT BINDING 654 654 HEME (COVALENT) (BY SIMILARITY).
 FT METAL 655 655 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
 SQ SEQUENCE 739 AA: 80944 MW: E681BB237ACB91F4 CRC64.

RL U. Biol. Chem. 268:12812-12817(1993).
 RN [6]
 RP 3D-STRUCTURE MODELING.
 RX MEDLINE-96128046; PubMed-8554505;
 RA Cozier G.E., Anthony C.;
 RT "Structure of the quinoprotein glucose dehydrogenase of *Escherichia coli* modelled on that of methanol dehydrogenase from *Methylobacterium extorquens*.";
 RL Biochem. J. 312:679-685(1995).
 CC -1- FUNCTION: GDH IS PROBABLY INVOLVED IN ENERGY CONSERVATION RATHER THAN IN SUGAR METABOLISM.
 CC -1- CATALYTIC ACTIVITY: D-glucose + acceptor = D-glucono-1,5-lactone + reduced acceptor.
 CC -1- COFACTOR: PO4.
 CC -1- SUBUNIT: MONOMER.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane.
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POQ DEHYDROGENASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X51323; CA35706.1; -
 DR EMBL: D12651; BA02174.1; -
 DR EMBL: D26562; CA820298.1; -
 DR EMBL: AE000122; AAC73235.1; -
 DR PIR: J01017; J01017.
 DR HSSP: P38539; 4AAH.
 DR Ecogene: EG10369; qcd.
 DR InterPro: IPR001479; Bac_POQ.
 DR InterPro: IPR002372; Bac_POQ_repeat.
 DR Pfam: PF01011; Bacterial_POQ_7.
 DR PROSITE: PS00363; BACTERIAL_POQ_1; 1.
 DR PROSITE: PS00364; BACTERIAL_POQ_2; 1.
 KW Oxidoreductase; POQ; Transmembrane; Inner membrane; Periplasmic; Complete proteome.
 FT DOMAIN 1 10
 FT TRANSMEM 11 37 CYTOPLASMIC (PROBABLE).
 FT DOMAIN 38 40 PERIPLASMIC (PROBABLE).
 FT TRANSMEM 41 58 PROBABLE.
 FT DOMAIN 59 62 CYTOPLASMIC (PROBABLE).
 FT TRANSMEM 63 81 PROBABLE.
 FT DOMAIN 82 95 PERIPLASMIC (PROBABLE).
 FT TRANSMEM 96 110 PROBABLE.
 FT DOMAIN 111 118 CYTOPLASMIC (PROBABLE).
 FT TRANSMEM 119 141 PROBABLE.
 FT DOMAIN 142 142 PERIPLASMIC (PROBABLE).
 FT ACT_SITE 466 466 BASE (PROBABLE).
 FT CONFLICT 59 59 R -> L (IN REF. 1).
 FT CONFLICT 149 156 TISADATP -> HIKRCHT (IN REF. 1 AND 2).
 FT CONFLICT 193 193 N -> K (IN REF. 1 AND 2).
 FT CONFLICT 666 666 Q -> H (IN REF. 2).
 SQ SEQUENCE 796 AA; 86747 MW; D9BDC705A12894E9 CRC64;

Query Match 25.6%; Score 127.5; DB 1; Length 796;
 Best Local Similarity 33.7%; Pred. No. 7,1e-06;
 Matches 35; Conservative 15; Mismatches 29; Indels 25; Gaps 5;

OY 6 LMAS-----AGATALLAPAFQAVTPVTELLANPPAGMISYQONQENRHSPLTQ 58
 DB 135 LTVAGNDPQEIINGTSLADATPAEA-ISPVAQ-----DWPAYKNGQGRSPKLO 185
 OY 59 ITTENYGOLOIYWA-----KGMOPGRV--QVTPLIHDGVYLI 93
 DB 186 INADNVHLKEAVFERTGDKOPNDPGEITNEVTPIKGVDLXL 229

RESULT 6

DHG_GLUOX
 ID DHG_GLUOX STANDARD: PRT; 808 AA.
 AC P27175;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Glucose dehydrogenase [pyrroloquinoline-quinone] precursor
 DE (EC 1.1.99.17).
 GN GDH.
 OS Gluconobacter oxydans (Gluconobacter suboxydans).
 OC Bacteria; Proteobacteria; alpha subdivision; Acetobacteraceae;
 OC Gluconobacter.
 OX NCBI_TaxID=442;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-92017653; PubMed-1833618;
 RA Cleton-Jansen A.-M., Dekker S., van de Putte P., Goosen N.;
 RT "A single amino acid substitution changes the substrate specificity of quinoprotein glucose dehydrogenase in *Gluconobacter oxydans*.";
 RL Mol. Gen. Genet. 229:206-212(1991).
 RN [2]
 RP REVISION TO 213.
 RA Goosen N.;
 RL Submitted (FEB-1992) to the EMBL/Genbank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: D-glucose + acceptor = D-glucono-1,5-lactone + reduced acceptor.
 CC -1- COFACTOR: PO4.
 CC -1- SUBCELLULAR LOCATION: OUTSIDE OF THE PERIPLASMIC MEMBRANE.
 CC -1- MISCELLANEOUS: THE P1 FORM CAN OXIDIZE ONLY D-GLUCOSE, WHILE THE P2 FORM CAN ALSO OXIDIZE MALTOSE. THE SEQUENCE OF P1 FORM IS SHOWN HERE.
 CC -----
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POQ DEHYDROGENASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X62710; CAA44594.1; ALT_SEQ.
 DR PIR: S17716; ORKEY.
 DR InterPro: IPR001479; Bac_POQ.
 DR InterPro: IPR002372; Bac_POQ_repeat.
 DR Pfam: PF01011; Bacterial_POQ_7.
 DR PROSITE: PS00363; BACTERIAL_POQ_1; 1.
 DR PROSITE: PS00364; BACTERIAL_POQ_2; 1.
 KW Oxidoreductase; POQ; Transmembrane; Periplasmic; Signal.
 FT SIGNAL 1 33
 FT CHAIN 34 808
 FT TRANSMEM 35 54 GLUCOSE DEHYDROGENASE [PYRROLOQUINOLINE-QUINONE].
 FT TRANSMEM 59 76 POTENTIAL.
 FT TRANSMEM 94 108 POTENTIAL.
 FT TRANSMEM 123 138 POTENTIAL.
 FT ACT_SITE 470 470 BASE (POTENTIAL).
 FT VARIANT 788 788 H -> N (IN P2 FORM).
 SQ SEQUENCE 808 AA; 87567 MW; 0F4160DA7852445 CRC64;

Query Match 21.7%; Score 108; DB 1; Length 808;
 Best Local Similarity 28.2%; Pred. No. 0.00084;
 Matches 29; Conservative 14; Mismatches 40; Indels 20; Gaps 3;

OY 13 ALATLAPAFQAVTPVTELLA-----NPPAGEWISYQONQENRHSPLTQIT 61
 DB 128 AVLALFASLFTPHISGLPTQIANSPADPDVTPASEMAYRTQAGDRWSPLANDINA 187
 OY 62 ENVGOLOIYWA-----ARGMOPGRV--QVTPLIHDGVYLI 95
 DB 188 TIVSNLKVAMHITKDMNNSNDPGEOTNEATPIEFNNITLYWCS 230

RESULT 7
DHML_METOR STANDARD: PRT: 626 AA.
ID DHML_METOR
AC P15279;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Methanol dehydrogenase subunit 1 precursor (EC 1.1.99.8) (MDH large alpha subunit) (MDH).
GN MOX.
OS Methylobacterium organophilum XX.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Methylobacterium group; Methylobacterium.
OX NCBI_TaxID=410;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 29-43.
RC STRAIN=ATCC 27886 / DSM 760 / NCIB 11278;
RX MEDLINE=89008094; PubMed=2459109;
RA Machin S.M., Hanson R.S.;
RT "Nucleotide sequence and transcriptional start site of the Methylobacterium organophilum XX methanol dehydrogenase structural gene";
RL J. Bacteriol. 170:4739-4747(1988).
CC -1- CATALYTIC ACTIVITY: A primary alcohol + acceptor = an aldehyde + reduced acceptor.
CC -1- COFACTOR: POO.
CC -1- SUBUNIT: THE HOLOENZYME MDH HAS AN ALPHA-2/BETA-2 CONFIGURATION OF A LARGE ALPHA AND A SMALL BETA SUBUNIT.
CC -1- SUBCELLULAR LOCATION: PERIPLASMIC, BUT ASSOCIATED WITH INNER MEMBRANE.
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POO DEHYDROGENASE FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC
CC EMBL: M22629; AAA50289.1; -.
CC HSSP: P38539; 4AAH.
DR InterPro: IPR001479; Bac_POO.
DR InterPro: IPR002372; Bac_POO_repeat.
DR Pfam: PF01011; Bacterial_POO_1; 1.
DR PROSITE: PS00363; BACTERIAL_POO_1; 1.
DR PROSITE: PS00364; BACTERIAL_POO_2; 1.
KW Oxidoreductase; POO; Signal; Methanol utilization; Periplasmic.
FT CHAIN 1 28 METHANOL DEHYDROGENASE SUBUNIT 1.
FT SIGNAL 29 626
FT DISULFID 130 131 BY SIMILARITY.
FT DISULFID 413 442 BY SIMILARITY.
FT ACT_SITE 330 330 BASE (POTENTIAL).
SQ SEQUENCE 626 AA; 68677 MW; 8768F6B8371E5DF CRC64;
Query Match 20.9%; Score 104; DB 1; Length 626;
Best Local Similarity 34.8%; Pred. No. 0.0017;
Matches 31; Conservative 18; Mismatches 34; Indels 6; Gaps 5;
OY 10 SAGALALA-AP-APAOVTPYTDLLANPAGE-WISGQNGENYRHSPLQTTEVVGQ 66
DB 7 SVSALAMLALAPALSSAVAYANDKLVLELSDDMVMGKKNYSNNYSELKQVKNKSVKQ 66
OY 67 LQLYW--ARGMQPKVQVYPLIHGDMVL 93
DB 67 LRPAMTFSTGLNGH-EGAPLVVDGKMTY 94
RESULT 8
DHGA_ACICA STANDARD: PRT: 801 AA.
ID DHGA_ACICA

AC P05465;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Glucose dehydrogenase-A [pyrroloquinoline-quinone] precursor (EC 1.1.99.17) (Quinoprotein glucose DH) (GDH-A).
GN GDHA.
OS Acinetobacter calcoaceticus.
OC Bacteria; Proteobacteria; gamma subdivision; Moraxellaceae;
OC Acinetobacter.
OX NCBI_TaxID=471;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LMD 79.41;
RX MEDLINE=88289368; PubMed=339393;
RA Cleton-Jansen A.M., Goosen N., Odle G., van de Putte P.;
RT "Nucleotide sequence of the gene coding for quinoprotein glucose dehydrogenase from Acinetobacter calcoaceticus";
RL Nucleic Acids Res. 16:6228-6228(1988).
CC -1- CATALYTIC ACTIVITY: D-glucose + acceptor = D-glucose-1,5-lactone + reduced acceptor.
CC -1- COFACTOR: POO.
CC -1- SUBUNIT: MONOMER.
CC -1- SUBCELLULAR LOCATION: Periplasmic.
CC -1- MISCELLANEOUS: ACINETOBACTER CALCOACETICUS CONTAINS TWO DIFFERENT POO DEPENDENT ENZYMES WITH GDH ACTIVITY. GDH-A PREFERS 2-DEOXYGLUCOSE AS SUBSTRATE, THE SPECIFIC SUBSTRATES FOR GDH-B ARE DISACCHARIDES.
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POO DEHYDROGENASE FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC
CC EMBL: X07235; CAA30222.1; -.
CC PIR: S00943; S00943.
DR InterPro: IPR001479; Bac_POO.
DR InterPro: IPR002372; Bac_POO_repeat.
DR Pfam: PF01011; Bacterial_POO_1; 1.
DR PROSITE: PS00363; BACTERIAL_POO_1; 1.
DR PROSITE: PS00364; BACTERIAL_POO_2; 1.
KW Oxidoreductase; POO; Transmembrane; Periplasmic; Signal.
FT CHAIN 1 33
FT SIGNAL 34 801
FT TRANSMEM 39 55 GLUCOSE DEHYDROGENASE-A
FT TRANSMEM 59 79 [PYRROLOQUINOLINE-QUINONE].
FT TRANSMEM 94 108 POTENTIAL.
FT TRANSMEM 119 138 POTENTIAL.
FT ACT_SITE 471 471 BASE (POTENTIAL).
SQ SEQUENCE 801 AA; 86956 MW; 2F4EA222EB5450D8 CRC64;
Query Match 19.1%; Score 95; DB 1; Length 801;
Best Local Similarity 31.2%; Pred. No. 0.02;
Matches 24; Conservative 11; Mismatches 22; Indels 20; Gaps 3;
OY 20 PAFACVTPYTDLLANPAGEWISGQNGENYRHSPLQTTEVVGQLOLYMARGMQPK 79
DB 152 PETAAVGVAE-----SDMPAYGRTQAGVRYSPKQINDQVKKLVAMT--LRTGD 202
OY 80 V-----QVPL 85
DB 203 LKTNDGSETTNOVTP 219
RESULT 9
DHML_METEX STANDARD: PRT: 626 AA.
ID DHML_METEX

AC P16027;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-APR-1990 (Rel. 14, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Methanol dehydrogenase subunit 1 precursor (EC 1.1.99.8) (MDH large
 alpha subunit) (MDH).
 GN MOXF.
 OS Methylobacterium extorquens.
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Methylobacterium group; Methylobacterium.
 OX NCBI_TaxID=408;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=AM1 / NCIB 9133;
 RX MEDLINE=90337342; PubMed=2116368;
 RA Anderson D.J., Morris C.J., Nunn D.N., Anthony C., Lidstrom M.E.;
 RT "Nucleotide sequence of the Methylobacterium extorquens AM1 moxf and
 RL gene 90:173-176(1990)."
 RN [2]
 RP SEQUENCE OF 28-53.
 RC STRAIN=AM1 / NCIB 9133;
 RX MEDLINE=89350892; PubMed=2504152;
 RA Nunn D.N., Day D., Anthony C.;
 RT "The second subunit of methanol dehydrogenase of Methylobacterium
 RL extorquens AM1".
 RN Biochem. J. 260:857-862(1989).
 RN [3]
 RP DISULFIDE BONDS.
 RX MEDLINE=95384759; PubMed=7656012;
 RA Blake C.C.F., Ghosh M., Harlos K., Avezoux A., Anthony C.;
 RT "The active site of methanol dehydrogenase contains a disulphide
 RL bridge between adjacent cysteine residues.";
 RN Nat. Struct. Biol. 1:102-105(1994).
 RN [4]
 RP X-RAY CRYSTALLOGRAPHY (1.94 ANGSTROMS).
 RX MEDLINE=95253818; PubMed=7735834;
 RA Ghosh M., Anthony C., Harlos K., Goodwin M.G., Blake C.;
 RT "The refined structure of the quinoprotein methanol dehydrogenase
 RL from Methylobacterium extorquens at 1.94 A.";
 RN Structure 3:177-187(1995).
 CC -1- CATALYTIC ACTIVITY: A primary alcohol + acceptor = an aldehyde +
 CC reduced acceptor.
 CC -1- COFACTOR: PQO.
 CC -1- SUBUNIT: THE HOLOENZYME MDH HAS AN ALPHA-2/BETA-2 CONFIGURATION
 CC OF A LARGE ALPHA AND A SMALL BETA SUBUNIT
 CC -1- SUBCELLULAR LOCATION: PERIPLASMIC, BUT ASSOCIATED WITH INNER
 CC MEMBRANE.
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL PQO DEHYDROGENASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M31108; AAA25380.1; -
 DR DR PIR: S07908; S07908.
 DR PIR: J00706; J00706.
 DR HSSP: P38539; AAAH.
 DR InterPro: IPR001479; Bac_PQO.
 DR InterPro: IPR002372; Bac_PQO_repeat.
 DR Pfam: PF01011; Bacterial_PQO; 7.
 DR PROSITE: PS00363; BACTERIAL_PQO_1; 1.
 DR PROSITE: PS00364; BACTERIAL_PQO_2; 1.
 KW Oxidoreductase; PQO; Signal; Methanol utilization; Periplasmic.
 FT SIGNAL 1 27
 FT CHAIN 28 626 METHANOL DEHYDROGENASE SUBUNIT 1.
 FT DISULFID 130 131
 FT DISULFID 413 442
 FT ACT_SITE 330 330 BASE (POTENTIAL).
 FT

SO SEQUENCE 626 AA; 68434 MW; 64988D0AFD2AD34C CRC64;
 Query Match 18.9%; Score 94; DB 1; Length 626;
 Best Local Similarity 34.8%; Pred. No. 0.019;
 Matches 31; Conservative 16; Mismatches 36; Indels 6; Gaps 5;
 QY 10 SAGALLALA-AP-APQVPTVDLLANPAGE-WISYQONQENYRSPQTTEWVG 66
 Db 7 SVSALMALALAPALSSGAYANDKLYELSKSDNNVMGKNYDNNFSDLKQINKGVAK 66
 QY 67 LQLYW--ARGMPKGVQVYPLIHGCVMTL 93
 Db 67 LRPAMFTSTGLNGH-EGAPLVVDKMYI 94
 RESULT 10
 DHM1_PARDE STANDARD; PRT; 631 AA.
 ID P12293;
 AC 01-OCT-1989 (Rel. 12, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Methanol dehydrogenase subunit 1 precursor (EC 1.1.99.8) (MDH large
 DE alpha subunit) (MDH).
 GN MOXF.
 OS Paracoccus denitrificans.
 OC Bacteria; Proteobacteria; alpha subdivision; Rhodospirillum group;
 OC Paracoccus.
 OX NCBI_TaxID=266;
 RN [1]
 RP SEQUENCE FROM N.A. AND SEQUENCE OF 33-49.
 RX MEDLINE=87307969; PubMed=3114231;
 RA Harms N., de Vries G.E., Maurer K., Hoogendijk J., Stoutamer A.H.;
 RT "Isolation and nucleotide sequence of the methanol dehydrogenase
 RL structural gene from Paracoccus denitrificans.";
 RN J. Bacteriol. 169:3969-3975(1987).
 CC -1- CATALYTIC ACTIVITY: A primary alcohol + acceptor = an aldehyde +
 CC reduced acceptor.
 CC -1- COFACTOR: PQO.
 CC -1- SUBCELLULAR LOCATION: Periplasmic.
 CC -1- MISCELLANEOUS: MDH IS THE MAJOR PROTEIN IN THE CELL DURING GROWTH
 CC ON METHANOL (IN P. DENITRIFICANS MDH CONSTITUTES UP TO 15% OF THE
 CC TOTAL CELL PROTEIN).
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL PQO DEHYDROGENASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M17339; AAA8366.1; -
 DR HSSP: P38539; AAAH.
 DR InterPro: IPR001479; Bac_PQO.
 DR InterPro: IPR002372; Bac_PQO_repeat.
 DR Pfam: PF01011; Bacterial_PQO; 7.
 DR PROSITE: PS00363; BACTERIAL_PQO_1; 1.
 DR PROSITE: PS00364; BACTERIAL_PQO_2; 1.
 KW Oxidoreductase; PQO; Signal; Methanol utilization; Periplasmic.
 FT SIGNAL 1 32
 FT CHAIN 33 631 METHANOL DEHYDROGENASE SUBUNIT 1.
 FT DISULFID 135 136 BY SIMILARITY.
 FT DISULFID 448 447
 FT ACT_SITE 335 335 BASE (POTENTIAL).
 FT SEQUENCE 631 AA; 69799 MW; 0934DC93FFC5730B CRC64;
 SO

Query Match 18.7%; Score 93; DB 1; Length 631;
 Best Local Similarity 35.8%; Pred. No. 0.025;
 Matches 34; Conservative 15; Mismatches 34; Indels 12; Gaps 6;

```

OY      4 TSLMASAGALAL--AAPAFQVTPVDEL--LAMPAGEWISYGOENYRHSPLTQIT 60
      :| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      12 SSIAMVAVMGLAVLTAPATA-----NDOLVELAKDPA-NMYMTGRDYNAGYSMTDIN 65
      :| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY      61 TENVGOLQLVW--ARGMGKQVQVPLIDHGYMYL 93
      :| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      66 KENVKOLRPAMSFSTGVLIHG-EGFTLVYGDRIIFI 99
      :| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 11
EXAA_PSEAE STANDARD; PRT; 623 AA.
ID EXAA_PSEAE
AC 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Quinoprotein ethanol dehydrogenase precursor (EC 1.1.99.-) (OEHD).
GN EXAA OR PA1982.
OS Pseudomonas aeruginosa.
OC Bacteria: Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OC NCBI_TaxID=287;
OX [1]
RN SEQUENCE FROM N.A.
RP STRAIN-ATCC 17933;
RC MEDLINE=99041560; PubMed=9826187;
RA Hickey M.J., Britkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA Gardner R.L., Goltz L., Tolentino E., Westbrock-Madman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
RA "Complete genome sequence of Pseudomonas aeruginosa PAO1, an
RA opportunistic pathogen.";
RL Nature 406:959-964(2000).
RN [3]
RP SEQUENCE OF 1-52 FROM N.A.
RC STRAIN-ATCC 17933;
RX MEDLINE=99173751; PubMed=10075429;
RA Schobert M., Goerisch H.;
RT "Cytochrome c550 is an essential component of the quinoprotein ethanol
RT oxidation system in Pseudomonas aeruginosa: cloning and sequencing of
RT the genes encoding cytochrome c550 and an adjacent acetaldehyde
RT dehydrogenase.";
RL Microbiology 145:471-481(1999).
RN [4]
RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF 35-616.
RC STRAIN-ATCC 17933;
RX MEDLINE=20202376; PubMed=10736230;
RA Kettel T., Diehl A., Knaute T., Stezowski J.J., Hoehne W.,
RA Goerisch H.;
RT "X-ray structure of the quinoprotein ethanol dehydrogenase from
RT Pseudomonas aeruginosa: basis of substrate specificity.";
RL J. Mol. Biol. 297:961-974(2000).
RN [1]
RP FUNCTION: OXIDIZE PRIMARY ALCOHOLS AND ALSO ACTS ON SECONDARY
RN ALCOHOL, BUT NOT HIGHLY ACTIVE ON METHANOL.
CC -1- COPFACTOR: POQ AND CALCIUM.
CC -1- SUBUNIT: HOMODIMER.
CC -1- SUBCELLULAR LOCATION: PERIPLASTIC.
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POQ DEHYDROGENASE FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

```

the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).

```

OY      3 PTSLMASAGALAL--LAPAFQVTPVDELALN--PAPGEMWISYGOENYRHSPLT 57
      :| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      9 PAGILRPSPILCHAFVAVGASAGALADYWMEDYANDKTTGDTVLQYGMGTHAORNSPLK 68
      :| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY      58 QITTEVVGOLQLVWARGM---OPGKQVQVPLIDHGYMYL 93
      :| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      69 QVNAVDFKLPAMSFSTGVLIHG-ESQAIYSDGVIIYV 107
      :| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 12
TNFA_MARMO STANDARD; PRT; 233 AA.
ID TNFA_MARMO
AC 035734;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Tumor necrosis factor precursor (TNF-alpha) (Cachectin).
GN TNF OR TNFA.
OS Marmota monax (Woodchuck).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Scuridae; Sciurinae;
OC Marmota.
OX NCBI_TaxID=9995;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Periplasmic blood;
RX MEDLINE=98139533; PubMed=9472070;
RA Lohrengel B., Lu W., Roggendorf M.;
RT "Molecular cloning of the woodchuck cytokines: TNF-alpha, IFN-gamma,
RT and IL-6.";
RL Immunogenetics 47:332-335(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE-Periplasmic blood;
RX Zhou H., Hu J., Seeger C.;
RT Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE-Periplasmic blood;
RX MEDLINE=20184748; PubMed=10721723;
RA Li D.H., Hayell E.A., Brown C.L., Cullen J.M.;
RT "Woodchuck lymphotoxin-alpha, -beta and tumor necrosis factor genes:
RT structure, characterization and biological activity.";
RL Gene 242:295-305(2000).
RN [1]
RP FUNCTION: TNF IS MAINLY SECRETED BY MACROPHAGES. IT IS A CYTOKINE
RN WITH A WIDE VARIETY OF FUNCTIONS: IT CAN CAUSE CYTOLYSIS OF

```

RESULT	13
PSAB_ODOSI	ID
AC	PSAB_ODOSI
AC	P49460;
DT	01-FEB-1996 (Rel. 33, Created)
DT	01-FEB-1996 (Rel. 33, Last sequence update)
DT	01-MAR-2002 (Rel. 41, Last annotation update)
DE	Photosystem I P700 chlorophyll A apoprotein A2 (Psab) (PSI-B).
GN	PSAB.
OS	Odentella sinensis.
OG	Chloroplast.
OC	Eukaryota; Stramenopiles; Bacillariophyta; Coscinodiscophyceae;
OC	Biddulphiophycidae; Eupodiscales; Eupodiscaceae; Odontella.
OX	NCBI_TaxID=2839;
RN	[1]
RP	SEQUENCE FROM N.A.
RA	Kowalik K.V., Stoebe B., Schaffran I., Kroth-Pancic P., Preler U.,

RESULT	14		
PSAB_HEATTR			
ID	PSAB_HEATTR	STANDARD;	PRT; 776 AA
AC	Q9X02;		
DT	01-MAR-2002	(Rel. 41, Created)	
DT	01-MAR-2002	(Rel. 41, Last sequence update)	

DT	01-MAR-2002 (Rel. 41, Last annotation update)
DE	Photosystem I P700 chlorophyll A apoprotein A2 (PsaB) (Psi-B).
GN	PSAB.
OS	Heterocapsa triquetra (Dinoflagellate).
OC	Chloroplast.
OC	Eukaryota; Alveolata; Dinophyceae; Peridinales; Heterocapsaceae;
OC	Heterocapsa.
OX	NCBI_TaxID=66468;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=CCMP 449;
RA	MEDLINE=99334925; PubMed=10408440;
RA	Zhang Z., Green B.R., Cavalier-Smith T.;
RT	"Single gene circles in dinoflagellate chloroplast genomes.";
RL	Nature 400:155-159(1999).
CC	-1- FUNCTION: PsaA and PsaB bind P700, the primary electron donor of photosystem I (PSI), as well as the electron acceptors A0, A1, and Fx. PSI functions as a plastocyanin/cytochrome c6-ferredoxin oxidoreductase.
CC	-1- COFACTOR: P700 is a chlorophyll A dimer, A0 is chlorophyll A, A1 is a phytylquinone and Fx is a 4Fe-4S iron-sulfur center.
CC	-1- SUBUNIT: A psaA/B heterodimer binds the P700 chlorophyll special pair and subsequent electron acceptors. The PSI reaction center of higher plants and algae is composed of one at least 11 subunits.
CC	-1- SUBCELLULAR LOCATION: Integral membrane protein. Chloroplast thylakoid membrane.
CC	-1- SIMILARITY: BELONGS TO THE PsaA/PsaB FAMILY.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or_send_an_email_to_license@isb-sib.ch).
CC	-----
DR	EMBL, AF130032; AAA04699.1; -
DR	InterPro: IPR001280; PsaA_PsaB.
DR	Pfam: PF00223; PsaA_PsaB; 2.
DR	PRINTS: PR00257; PROTSYSPSAB.
DR	PROSITE: PS00419; PHOTOSYSTEM_I_PsaAB; 1.
KM	Photosynthesis; Photosystem I; Electron transport; Chloroplast;
KM	Thylakoid; Transmembrane; Iron-sulfur; 4Fe-4S; Chlorophyll.
KM	TRANSMEM 57 80
FT	TRANSMEM 150 172
FT	TRANSMEM 233 257
FT	TRANSMEM 334 352
FT	TRANSMEM 382 405
FT	TRANSMEM 421 447
FT	TRANSMEM 472 494
FT	TRANSMEM 561 579
FT	TRANSMEM 619 640
FT	TRANSMEM 687 709
FT	TRANSMEM 753 773
FT	METAL 603 603
FT	METAL 612 612
FT	BINDING 698 698
FT	BINDING 706 706
FT	BINDING 714 714
FT	BINDING 715 715
FT	BINDING 746 746
FT	SEQUENCE 776 AA; 68143 MW; F8640CA5F58C913 CRC64;

	Query Match	Best Local Similarity	Matches	Score	Pred. Mismatches	DB 1	Length	Indels	Gaps
QY	4	14.68;	28;	72.5;	18;		776;		
	TSLLMSAGALALALAA	25.98;			39;				
	-----FAQVTPYDEL								
	LAAP-----AGEMIS								
	YTGQENY						51		
Db	244	TSLLMSAGALALALAA	25.98;	72.5;	18;	39;	776;	23;	6;
	AGLHIVALPASRGIVHT								
	STSPVDSFTFAPSKAL								
	PFYSGWAVYAAQDID								
	DD								
	301								

Oy 52 RHSPLETTENVGOLVNRMG--PGKYVPPLHD---GYMYL 93
| : | : | : | : | : | :
Db 302 NH---TFSTVGAGCAITFLGCVNSDASLYLFDIAHHHLAIGVLFI 346

RESULT	15			
AC	XOXF_PARDE	STANDARD;	PRI:	600 AA.
AC	P29968;			
DT	01-APR-1993 (Rel. 25, Created)			
DT	01-OCT-1996 (Rel. 34, Last sequence update)			
DT	15-DEC-1998 (Rel. 37, Last annotation update)			
DE	Putative dehydrogenase XOXF precursor (EC 1.1.99.-).			
GN	XOXF.			
OS	Paracoccus denitrificans.			
OC	Bacteria; Proteobacteria; alpha subdivision; Rhodobacter group;			
OC	Paracoccus.			
OX	NCBI_TaxID=266;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Harms N.;			
RL	Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.			
RP	[2]			
RC	SEQUENCE OF 380-600 FROM N.A.			
RC	STRAIN-PD 1235;			
RC	MEDLINE=92041583; PubMed=1657873;			
RA	Ras J., Reijnders W.N.M., van Spanning R.J.M., Harms N., Oltmann L.F.,			
RA	Stouthamer A.H.;			
RT	"Isolation, sequencing, and mutagenesis of the gene encoding			
RT	cytochrome c531 of Paracoccus denitrificans and characterization of			
RL	the mutant strain";			
RL	J. Bacteriol. 173:6971-6979(1991).			
CC	-1- COPACITOR: POQ (BY SIMILARITY).			
CC	-1- SIMILARITY: BELONGS TO THE BACTERIAL POQ DEHYDROGENASE FAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation - its			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by, and for commercial			
CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL: U13436; AAC44555.1; -			
DR	EMBL: M75583; AAA25574.1; -			
DR	PIR: A41378; A41378.			
DR	HSSP: P38539; AAAH.			
DR	InterPro: IPR002372; Bac_P0Q.repeat.			
DR	Pfam: PF01011; Bacterial_POQ: 7.			
KW	Oxidoreductase; POQ; Signal.			
FT	SIGNAL	1	21	
FT	CHAIN	22	600	POTENTIAL.
FT	ACT_SITE	318	318	PUTATIVE DEHYDROGENASE XOXF.
FT	SEQUENCE	600 AA: 65159 MW; DCA99661BCC5A3CE CRC64;		BASE (POTENTIAL).

	Query Match	14.2%	Score 70.5;	DB 1;	Length 600;
	Best Local Similarity	29.0%;	Pred. No. 5.7;		
	Matches 27; Conservative	14;	Mismatches 33;	Indels 19;	Gaps 4;
QY	10 SAGALALLAPAFAYQVPTVDLLANP-----PAGEWISTGQCNENRHSPLQITTE	62			
	: : : : :				
Db	6 NGACIALMLMSGTAA-----LANEORAGRDRCAPOAWIGDYANTRYSLTDINKD	56			
QY	63 NVCGLQLWV--ARGMQPGKYQVTPFLHDCMYL	93			
	: : : : : :				
Db	57 NVKDLRVAMTFSTGVLRGH-EGSPILVGDVMY	88			

Search completed: May 24, 2002, 10:15:47
Job time: 765 sec

Fri May 24 11:27:23 2002

us-08-934-506a-5_copy_1_95.rsp

Fri May 24 11:27:24 2002

us-08-934-506a-5_copy_1_95.rspt

Page 1

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2002, 10:05:07 ; Search time 93.93 Seconds
(without alignments)
174.966 Million cell updates/sec

Title: US-08-934-506a-5_COPY_1_95
Perfect score: 498
Sequence: 1 MKPSTLMASGALALAAP.....QPKGVVTPPLIHGVYLAN 95

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: SPTRMBL_19:*
2: SP_archaea:*
3: SP_bacteria:*
4: SP_fungi:*
5: SP_human:*
6: SP_invertebrate:*
7: SP_mammal:*
8: SP_mhc:*
9: SP_organelle:*
10: SP_plant:*
11: SP_todent:*
12: SP_virus:*
13: SP_vertebrate:*
14: SP_unclassified:*
15: SP_rvirus:*
16: SP_bacteriap:*
17: SP_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	157.5	31.6	608	2	Q93RE9	Q93RE9 pseudogluc
2	139.5	28.0	742	2	Q53362	Q53362 acetobacter
3	130.5	26.2	691	2	Q9AF95	Q9AF95 pseudomonas
4	130	26.1	695	2	P95466	P95466 pantoea cit
5	128	25.7	695	2	Q9P9U2	Q9P9U2 pseudomonas
6	127	25.5	623	2	Q9ACW3	Q9ACW3 pseudomonas
7	121	24.3	785	16	Q9KFF6	Q9KFF6 rhizobium 1
8	118.5	23.8	708	2	Q46444	Q46444 comonomas t
9	113.5	22.8	777	16	Q92R83	Q92R83 rhizobium m
10	108.5	21.8	803	16	Q91115	Q91115 pseudomonas
11	107.5	21.6	698	2	Q9KH03	Q9KH03 alcaligenes
12	103	20.7	629	2	Q9A048	Q9A048 methylolact
13	99	19.9	601	2	Q9EYR8	Q9EYR8 rhizobium m
14	99	19.9	601	16	Q92MY9	Q92MY9 rhizobium m
15	92	18.5	633	2	Q24759	Q24759 hypomicrob
16	90	18.1	790	2	Q9X255	Q9X255 pantoea cit

17	87.5	17.6	601	2	P71509	P71509 methylolact
18	83.5	16.8	695	2	Q934G0	Q934G0 pseudomonas
19	73.5	14.8	1229	5	Q9VM05	Q9VM05 drosophila
20	72.5	14.6	306	16	Q9RY72	Q9RY72 deinococcus
21	70	14.1	407	16	Q9ACE2	Q9ACE2 caulobacter
22	70	14.1	877	4	Q9H3Q8	Q9H3Q8 homo sapien
23	70	14.1	878	4	Q9H3Q7	Q9H3Q7 homo sapien
24	70	14.1	957	4	Q14651	Q14651 homo sapien
25	69.5	14.0	1217	4	Q9UKW9	Q9UKW9 homo sapien
26	69.5	14.0	1234	5	Q24690	Q24690 drosophila
27	69	13.9	711	5	Q9YGD8	Q9YGD8 drosophila
28	68.5	13.8	447	16	Q9RUD4	Q9RUD4 deinococcus
29	68	13.7	223	3	Q99015	Q99015 trichoderma
30	68	13.7	599	2	Q91935	Q91935 methyllovora
31	68	13.7	856	16	Q9A6E5	Q9A6E5 caulobacter
32	68	13.7	878	4	Q9G222	Q9G222 homo sapien
33	68	13.7	901	4	Q9H195	Q9H195 homo sapien
34	68	13.7	1057	11	Q9EON4	Q9EON4 mus musculus
35	67.5	13.6	1085	5	Q24363	Q24363 drosophila
36	67	13.5	197	16	Q9A829	Q9A829 caulobacter
37	67	13.5	738	2	Q70031	Q70031 streptomyces
38	66.5	13.4	428	2	Q9LH85	Q9LH85 thiolactillu
39	66.5	13.4	647	16	Q9HUB1	Q9HUB1 pseudomonas
40	66.5	13.4	1516	16	Q9HVT2	Q9HVT2 pseudomonas
41	66	13.3	465	5	Q18914	Q18914 caenorhabdi
42	66	13.3	1017	5	Q9XVY2	Q9XVY2 lymantria d
43	66	13.3	1419	5	Q9YVL3	Q9YVL3 drosophila
44	66	13.3	1668	5	Q76930	Q76930 drosophila
45	65.5	13.2	1320	11	Q9UK25	Q9UK25 rattus norv

ALIGNMENTS

RESULT	ID	Q93RE9	PRELIMINARY:	PRT:	608 AA.
AC	Q93RE9	01-DEC-2001 (TREMBLrel. 19, Created)			
DT	01-DEC-2001 (TREMBLrel. 19, Last sequence update)				
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)				
DE	ALCOHOL DEHYDROGENASE.				
GN	ADH.				
OS	Pseudoglucanobacter saccharotogenes.				
OC	Bacteria; Pseudoglucanobacter.				
OX	NCBI_TaxID=133921;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=IFO 14464;				
RC	Shibata T., Saito Y.;				
RT	"Alcohol dehydrogenase."				
RL	Submitted (JUL-2000) to the EMBL/Genbank/DBJ databases.				
DR	EMBL: AB046580; BAB62258.1;				
SQ	SEQUENCE 608 AA; 65101 MW; 0ACECG7AE11BA570 CRC64;				

Query Match	31.6%	Score: 157.5;	DB 2;	Length 608;
Best Local Similarity	34.3%	Pred. No. 1.1e-08;		
Matches	36;	Conservative	15;	Mismatches 35;
				Indels 19;
				Gaps 1;
OY	10	SAGALALAAPAFNAO-----VPPVDELLANPFAEMISYGNQEN 50		
DB	16	STALIASLSPAFQHDHANAEPKAGSAIENFQVPTADLAGKNPANPILRGNTYG 75		
OY	51	YRHSPLTOITTEANGOLQVWARGMOPKQVTPPLIHGVYLAN 95		
DB	76	WGSPLDQINDKNDVGDLOLVMSRTMEPGSNAGAIATNGVIFLGN 120		
RESULT	2			
ID	Q53362	PRELIMINARY:	PRT:	742 AA.
AC	Q53362; Q44159;			

DT 01-NOV-1996 (Tremblrel. 01, Created)
 DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
 DE DEHYDROGENASE SUBUNIT OF ALCOHOL DEHYDROGENASE.
 OS Acetobacter pasteurianus (Acetobacter turbidans).
 OC Bacteria; Proteobacteria; alpha subdivision; Acetobacteraceae;
 NCBI_TaxID=438;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NC11380;
 RX MEDLINE=94042848; PubMed=8226628;
 RA Takemura H., Konô K., Horinouchi S., Beppu T.;
 RT "Induction by ethanol of alcohol dehydrogenase activity in Acetobacter
 pasteurianus.";
 RL J. Bacteriol. 175:6857-6866(1993).
 DR EMBL; D13893; BAA0252.1; F1G.
 DR HSP; Q924J7; 1F1G.
 DR InterPro: IPR001479; Bac_PQO.
 DR InterPro: IPR002372; Bac_PQO_repeat.
 DR Pfam: PF01011; Bacterial_PQO_6.
 DR PROSITE; PS00363; BACTERIAL_PQO_1; 1.
 DR PROSITE; PS00364; BACTERIAL_PQO_2; 1.
 SQ SEQUENCE 742 AA; 81680 MW; 031C4E9B6B13EC98 CRC64;

Query Match 28.0%; Score 139.5; DB 2; Length 742;
 Best Local Similarity 40.0%; Pred. No. 1.2e-06;
 Matches 34; Conservative 10; Mismatches 38; Indels 3; Gaps 2;

QY 11 AGALALAAPAFQVTVTEDEL-ANPAGSEISYSGONENRHSPLQITTEVNGOLQ 68
 DB 23 AALPYAAVPAADGGGNTGEEIHHDDHNNKSTGRYSQRRSPLOINRNVGDLK 82
 QY 69 LVNARGNQPKVQ-VPLIHDCVY 92
 DB 83 LAMVITIDTNRGGATPPLVDGIMY 107

RESULT 3
 ID 09AF95 PRELIMINARY; PRT; 691 AA.
 AC 09AF95;
 DT 01-JUN-2001 (Tremblrel. 17, Created)
 DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE 1-BUTANOL DEHYDROGENASE BDH.
 OS Pseudomonas butanovora.
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 NCBI_TaxID=86174;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Vangnal A.S., Arp D.J., Sayavedra-Soto L.A.;
 RT "Characterization of the expression of two distinct alcohol
 dehydrogenases involved in butane metabolism in Pseudomonas
 butanovora.";
 RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF355798; AAK27220.2;
 SQ SEQUENCE 691 AA; 75070 MW; 4FC7ED20CDA14P64 CRC64;

Query Match 26.2%; Score 130.5; DB 2; Length 691;
 Best Local Similarity 42.9%; Pred. No. 1e-05;
 Matches 24; Conservative 14; Mismatches 17; Indels 1; Gaps 1;

QY 39 GEMTISYSGONENRHSPLQITTEVNGOLQLVNARGMOPGK-VQVPLIHDCVY 93
 DB 41 GEMTISYSGONENRHSPLQITTEVNGOLQLVNARGMOPGK-VQVPLIHDCVY 96

RESULT 4
 P95466

ID P95466 PRELIMINARY; PRT; 786 AA.
 AC P95466;
 DT 01-MAY-1997 (Tremblrel. 03, Created)
 DT 01-MAY-1997 (Tremblrel. 03, Last sequence update)
 DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
 DE GLUCOSE DEHYDROGENASE [PYROLOQUINOLINE-QUINONE] (EC 1.1.99.17).
 GN GDH.
 OS Pantoea citrea.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 NCBI_TaxID=53336;
 RN [1]
 RP SEQUENCE FROM N.A. AND ACTIVITY.
 RC STRAIN=1056R;
 RX MEDLINE=97133947; PubMed=8979341;
 RA Cha J.-S., Pujol C., Kado C.I.;
 RT "Identification and characterization of a Pantoea citrea gene encoding
 glucose dehydrogenase that is essential for causing pink disease of
 pineapple.";
 RL Appl. Environ. Microbiol. 63:71-76(1997).
 CC -1- CATALYTIC ACTIVITY: D-GLUCOSE + ACCEPTOR -> D-GLUCONO-1,5-LACTONE +
 REDUCED ACCEPTOR.
 CC -1- COFACTOR: PQO.
 CC -1- SUBCELLULAR LOCATION: PERIPLASMIC (BY SIMILARITY).
 CC -1- MISCELLANEOUS: RESPONSIBLE FOR PINK TO BROWN DISCOLORATIONS OF
 INFECTED PINEAPPLES UPON HEATING DURING THE CANNING PROCEDURE.
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL PQO DEHYDROGENASE FAMILY.
 DR EMBL; X95985; CAA5229.1; -
 DR InterPro: IPR001479; Bac_PQO.
 DR InterPro: IPR002372; Bac_PQO_repeat.
 DR Pfam: PF01011; Bacterial_PQO_6.
 DR PROSITE; PS00364; BACTERIAL_PQO_2; 1.
 KW Oxidoreductase; PQO; Transmembrane; Cytoplasmic; Periplasmic.
 FT DOMAIN 1 13
 FT TRANSMEM 14 55
 FT DOMAIN 56 59
 FT TRANSMEM 60 80
 FT DOMAIN 81 85
 FT TRANSMEM 86 106
 FT DOMAIN 107 115
 FT TRANSMEM 116 136
 FT DOMAIN 137 431
 FT TRANSMEM 432 452
 FT DOMAIN 453 652
 FT TRANSMEM 653 673
 FT DOMAIN 674 700
 FT TRANSMEM 701 721
 FT DOMAIN 722 786
 FT TRANSMEM 770 770
 FT DOMAIN 770 770
 FT ACT_SITE 462 462
 FT ACT_SITE 786 786
 SQ SEQUENCE 786 AA; 86038 MW; 372402AAD8067CC CRC64;

Query Match 26.1%; Score 130; DB 2; Length 786;
 Best Local Similarity 34.2%; Pred. No. 1.4e-05;
 Matches 35; Conservative 14; Mismatches 35; Indels 26; Gaps 5;

QY 2 KPTSLWASGALALAAPAFQ-----VPTVTEDELANP-----AGETISYSGONQ 48
 DB 115 KPLAWITIGINALLIGSLHDPQINGVLVNSDK-----PPASASASADWPAYGRTO 170
 QY 49 ENRHSPLQITTEVNGOLQLVN-----ARGMOPGK-VQVPLIHDCVY 93
 DB 171 EGKYSPLQITTEVNGOLQLVN-----ARGMOPGK-VQVPLIHDCVY 96

RESULT 5
 ID 09F902 PRELIMINARY; PRT; 695 AA.
 AC 09F902;
 DT 01-MAR-2001 (Tremblrel. 16, Created)
 DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)

Query Match	25.7%	Score 128;	DB 2;	Length 695;
Best Local Similarity	38.3%	Pred. No. 1,9e-05;		
Matches	31;	Conservative	14;	Mismatches 32; Indels 4; Gaps 3;

QY	15	ALAAPFAQVTPYVDLL--ANPAGEMISGQNGENYRHPITQTTTENGSLQLYWA	72
		: : : :	
Db	16	ALIVA-AGAAAKAYDDAARASEQDGSLSHGRTYAEQRSPPLKQIDAGNVKGLGANY	74
QY	73	RGMPGK-VQVTEPLIHGYMY	92
		: : : : : :	
Db	75	LDLENNRGLEATPLVSDGVLY	95

RESULT	6	
Q9AGW3		
ID	Q9AGW3	PRELIMINARY; PRT: 623 AA.
AC	Q9AGW3;	
DT	01-JUN-2001 (TREMBLrel. 17, Created)	
DT	01-JUN-2001 (TREMBLrel. 17, Last sequence update)	
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)	
DE	1-BUTANOL DEHYDROGENASE BOH PRECURSOR.	
OS	<i>Pseudomonas butanovora</i> .	
OC	Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;	
OC	<i>Pseudomonas</i> .	
OX	NCBI_TaxID=86174;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RA	Sayavedra-Soto L.A., Vangnal A.S., Arp D.J.;	
RT	"Characterization of the expression of two distinct alcohol	
RT	dehydrogenases involved in butane metabolism in <i>Pseudomonas</i>	
RT	<i>butanovora</i> ."	
RL	Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.	
RL	EMBL: AF326086; AAK15506.1; "	
DR	HSSP: Q924J7; 1F6G.	
DR	InterPro: IPR002372; Bac_PQQ_repeat.	
DR	Pfam: PF01011; Bacterial_PQQ_4.	
KW	Signal.	
FT	SIGNAL	POTENTIAL.
FT	SEQUENCE	623 AA: 67553 MW: A572A1C64AC12D55 CR664.

```

Query Match      25.5%;   Score 127;   DB 2;   Length 623;
Best Local Similarity 38.6%;   Pred. No. 2,2e-05;
Matches 39; Conservative 12; Mismatches 34; Indels 16; Gaps 5;

2 KPSILMASGALALAAAPFAQVPTVDELLAN--PPAGEWISYGONEDNTHSPLETOI 59
  1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1

```

Query Match	24.3%	Score 121	DB 16	Length 785	
Best Local Similarity	31.5%	Prod. No. 0.00013			
Matches	39	Conservative	10	Mismatches	35
				Indels	40
				Gaps	4
QY	8	WASG-----	ALALLAAPAFQVTPVTD-----	ELLANPAGE	40
			: : : :		
Db	111	WARRGLAGPDGAPRLAVLAVAGSMTADPFKIDGALDTPDKVIRKANLINDNPAGE			170
QY	41	WISYQNGENYRHSPLTQITTEWQGLQLVARCMQPKY-----	QVPLIHDS		89
			: : : : : :		
Db	171	WHYHGRTQFGORSPLDQITTPDNVANLPAMT--YRTGDVAGPDIDGETTYQVTPPLKIGE			228
QY	90	VMYL	93		
		: :			
Db	229	TLXI	232		

RESULT	8		
ID	046444	PRELIMINARY;	PRT; 708 AA.
AC	046444;		
DT	01-NOV-1998 (TREMblrel. 08, Created)		
DT	01-NOV-1998 (TREMblrel. 08, last sequence update)		
DT	01-DEC-2001 (TREMblrel. 19, last annotation update)		
DE	QUINOAEMOPROTEIN ETHANOL DEHYDROGENASE TYPE I PRECURSOR (EC 1.1.99.-) (OH-EDHL).		
DE	QHEDH.		
GN	Comamonas testosteroni (Pseudomonas testosteroni).		
OC	Comamonas testosteroni; beta subdivision; Comamonadaceae; Comamonas.		
OC	Bacteria; Proteobacteria;		
OX	NCBI_TaxID=285;		
RN	[1]		
RN	SEQUENCE FROM N.A.		
RC	STRAIN-ATCC 15667;		
RX	MEDLINE=96184549; PubMed=8654419;		

RA Stoorvogel J., Kraayveld D.E., van Stuls C.A., Jongejan J.A.,
 RA De Vries S., Duine J.A.;
 RT "Characterization of the gene encoding quinohemoprotein ethanol
 RT dehydrogenase of Comamonas testosteroni";
 RL Eur. J. Biochem. 235:690-698(1996).
 RN [2]
 RP SEQUENCE OF 32-54 AND 477-490.
 RC STRAIN-ATCC 15667;
 RX MEDLINE-95324580; PubMed-7601151;
 RA De Jong G.A.H., Geerloff A., Stoorvogel J., Jongejan J.A., De Vries S.,
 RA Duine J.A.;
 RT "Quinohemoprotein ethanol dehydrogenase from Comamonas testosteroni.
 RT Purification, characterization, and reconstruction of the apoenzyme
 RT with pyruvate analogues";
 RL Eur. J. Biochem. 230:899-905(1996).
 RN [3]
 RP CHARACTERIZATION.
 RX MEDLINE-86242113; PubMed-3521592;
 RA Groen B.W., van Kleef M.A., Duine J.A.;
 RT "Quinohemoprotein alcohol dehydrogenase apoenzyme from Pseudomonas
 RT testosteroni";
 RL Biochem. J. 234:611-615(1986).
 RN [4]
 RP CRYSTALLIZATION.
 RX MEDLINE-21536088; PubMed-11679760;
 RA Oubrie A., Huizinga E.G., Roseboom H.J., Kalk K.H., de Jong G.A.H.,
 RA Duine J.A., Dijkstra B.W.;
 RT "Crystallization of quinohemoprotein alcohol dehydrogenase from
 RT Comamonas testosteroni: crystals with unique optical properties";
 RL Acta Crystallogr. D 57:1732-1734(2001).
 CC -1- FUNCTION: CATALYZES THE DYE-LINKED OXIDATION OF PRIMARY ALCOHOLS
 CC TO THE CORRESPONDING ALDEHYDES AND THE (SUBSEQUENT) OXIDATION OF
 CC THE ALDEHYDES TO CARBOXYLIC ACIDS.
 CC -1- COFACTOR: POO, HEME, AND CALCIUM.
 CC -1- SUBUNIT: MONOMER.
 CC -1- SUBCELLULAR LOCATION: PERIPLASMIC (POTENTIAL).
 CC -1- INDUCTION: BY ETHANOL AND BUTANOL.
 CC -1- SIMILARITY: TO ALCOHOL DEHYDROGENASE (ACCEPTOR) PROTEINS.
 DR EMBL: X81880; CAA57464.1; -
 DR HSSP: Q9447; JFLG.
 DR InterPro: IPR002372; Bac_POO_repeat.
 DR InterPro: IPR003088; Cyt_C1.
 DR InterPro: IPR002329; Cyt_C1C.
 DR Pfam: PF01011; Bacterial_POO; 6.
 DR Pfam: PF00034; cytochrome_c; 1.
 DR PRINTS: PR00605; CYCHROME.C1.
 KW Signal; POO; Heme; Calcium; Oxidoreductase; Periplasmic.
 FT SIGNAL 1 31
 FT CHAIN 32 708
 FT BINDING 635 635 QUINOHAEOPROTEIN ETHANOL DEHYDROGENASE
 FT BINDING 638 638 HEME (COVALENT) (BY SIMILARITY).
 FT METAL 639 639 HEME (COVALENT) (BY SIMILARITY).
 FT METAL 639 639 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
 SQ SEQUENCE 708 AA; 76822 MW; 99AB54BD6ACCA3 CRC64;

Query Match 23.8%, Score 118.5; DB 2; Length 708;
 Best Local Similarity 30.5%; Pred. No. 0.00021;
 Matches 32; Conservative 18; Mismatches 40; Indels 15; Gaps 4;

OY 3 PSLMASAGALALAAFAFAVPTV-----DELANP-AGWISYGOEN 50
 DB 12 PGRMWLAACLG--SAAFAQTGPAAOAAAVGVVDGDFIRANMARTPMDPTGVDAE 69
 OY 51 YRHSPLQITTEVGOLOLVARMOGPK-VQVPLIHDSVMTLA 94
 DB 70 TRYSLDOINANVKDIGLANSYMLSTREVTPEVVDGIMTVS 114

RESULT 9
 ID Q92RB3 PRELIMINARY; PRT; 777 AA.
 AC Q92RB3;

DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE PROBABLE GLUCOSE DEHYDROGENASE (PYROLODOUINOLINE-QUINONE) PROTEIN
 DE (EC 1.1.99.17).
 OS Rhizobium meliloti (Sinorhizobium meliloti).
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Rhizobiaceae; Sinorhizobium.
 NCBI_TaxID=382;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=1021.
 RX MEDLINE-21368234; PubMed-11474104;
 RA Galibert F., Flan T.M., Long S.R., Puehler A., Abola P., Ampe F.,
 RA Bailly-Hubler F., Barnett M.J., Becker A., Boistard P., Bothe G.,
 RA Boutry M., Bowser L., Buhrmester J., Cadieu E., Capela D., Chain P.,
 RA Coule A., Davis R.W., Dreano S., Federspiel N.A., Fisher R.F.,
 RA Gloux S., Godrie T., Goffeau A., Golding B., Gouzy J., Gurjel M.,
 RA Hernandez-Lucas I., Hong A., Huizar L., Hyman R.W., Jones T., Kahn D.,
 RA Kahn M.L., Kalman S., Keating D.H., Kiss E., Komp C., Lelaure V.,
 RA Masuy D., Palm C., Peck M.C., Pohl T.M., Portetelle D., Purnelle B.,
 RA Rampsperger U., Surzycki R., Thebaud P., Vandebol M.,
 RA Vorholter F.J., Weidner S., Thebaud P., Wells D.H., Wong K., Yeh K.-C., Batut J.;
 RT "The composite genome of the legume symbiont Sinorhizobium meliloti";
 RL Science 293:668-672(2001).
 DR EMBL: AL591785; CAC4553.1; -
 KW Oxidoreductase; Complete proteome.
 SQ SEQUENCE 777 AA; 84174 MW; D2F478A38BB10430 CRC64;

Query Match 22.8%; Score 113.5; DB 16; Length 777;
 Best Local Similarity 40.0%; Pred. No. 0.00082;
 Matches 28; Conservative 5; Mismatches 24; Indels 13; Gaps 2;

OY 35 NPRAGEWISGQENYRHSPLQITTEVGOLOLVARMOGPKV-----OVT 83
 DB 159 NVPGEHMOGRFGORSPDIDIPENVAMLVAVMO--VOTGVKLPEDVSETTYQVT 216
 OY 84 PLIHDSVMTLA 93
 DB 217 PLKVKDTLYV 226

RESULT 10
 ID Q91115 PRELIMINARY; PRT; 803 AA.
 AC Q91115;
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-OCT-2001 (TREMBLrel. 18, Last annotation update)
 DE GLUCOSE DEHYDROGENASE.
 GN GCD OR PA2290.
 OS Pseudomonas aeruginosa.
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 OC Pseudomonas.
 NCBI_TaxID=287;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ATCC 15692 / PA01;
 RX MEDLINE-20437337; PubMed-10984043;
 RA Stover C.K., Pham X.-O.T., Erwin A.L., Mizoguchi S.D., Warrenner P.,
 RA Hickey M.J., Brinkman F.S.L., Huftagle W.O., Kowalik D.J., Lagrou M.,
 RA Gader R.L., Goltzy L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
 RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larig K., Lim R.M.,
 RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
 RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
 RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
 RT opportunistic pathogen";
 RL Nature 406:959-964(2000).
 DR EMBL: AE004654; AAG05678.1; -
 DR InterPro: IPR002372; Bac_POO_repeat.
 DR Pfam: PF01011; Bacterial_POO; 7.
 KW Complete proteome.

SEQUENCE 803 AA; 86217 MW; 70DF46B0FE5ED249 CRC64;

Query Match
Best Local Similarity 27.5%; Pred. No. 0.003; Length 803;
Matches 36; Conservative 14; Mismatches 34; Indels 47; Gaps 5;

OY 6 LHM-----ASAGALALAAPAFACVPTVDLAMP-----36
DB 103 LHMPLPFRRLADGPAFLGTAALGVAAYLAGAAVGSQ-FTNPGQIVGRIDRSGMTST 161
OY 37 ----PAGEMISTGONCENTRHSPLQITTEENGQOLVWARMQSGKV-----Q 81
DB 162 APAMPDGMQAYGRTEFFGDRYSPLQITPANVQLEAMR--IRTGDLPTADDPLELINE 219
OY 82 VPPILHDGVMY 92
DB 220 NTPILKVMGLY 230

RESULT 11

O9KH03 PRELIMINARY; PRT; 698 AA.

AC O9KH03: 01-OCT-2000 (TREMblrel. 15, Created)
DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
DE TERAYDROFURFURYL ALCOHOL DEHYDROGENASE.
OS Alcaligenes eutrophus (Ralstonia eutropha).
OC Bacteria; Proteobacteria; beta subdivision; Ralstonia group;
OC Ralstonia.
NCBI_TaxID=510;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21125557; PubMed=11222593;
RA Zarrat G., Schrader T., Andreessen J.R.;
RT Catalytic and Molecular Properties of the Quinohemoprotein
Tetrahydrofuryl Alcohol Dehydrogenase from Ralstonia eutropha
RT Strain Bo.1.
RL J. Bacteriol. 183:1954-1960(2001).
DR EMBL; AF277373; AAF86335.1; --
DR HSSP; Q924J7; 1FLG.
DR InterPro: IPR002372; Bac_PQQ_repeat.
DR InterPro: IPR000345; CytC_heme_bind.
DR InterPro: IPR003088; Cyt.CI.
DR Pfam; PF01011; Bacterial_PQQ_6.
DR Pfam; PF00034; Cytochrome_c_1.
DR PROSITE; PS00190; CYTOCHROME_C; UNKNOMN_1.
SQ SEQUENCE 698 AA; 75471 MW; 9C87F180DBID2FB2 CRC64;

Query Match 21.6%; Score 107.5; DB 2; Length 698;
Best Local Similarity 34.4%; Pred. No. 0.003; Indels 5; Gaps 3;
Matches 31; Conservative 14; Mismatches 40;

OY 10 SAGALALAAPAFACVPTVDLAMPAG--EWISYSGONCENTRHSPLQITTEENG 65
DB 14 AASVALPAPAFAGANAARVDGAATIRANRGTNMPSTGYDVAETRFSLKLEOVNAGNVR 73
OY 66 QLOLVWARMQSGKV-VQVTPILHDGVMYLA 94
DB 74 NIGLAWSYDLESTRGVEATPLVVGVMYVS 103

RESULT 12

O9AQ48 PRELIMINARY; PRT; 629 AA.

AC O9AQ48: 01-JUN-2001 (TREMblrel. 17, Created)
DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)
DE METHANOL DEHYDROGENASE MXAF.
GN MXAF.

OS Methylobacterium nodulans.

OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;

OC Methylobacterium group; Methylobacterium.

NCBI_TaxID=114616;

RN [1]
RP SEQUENCE FROM N.A.

RC STRAIN=ORS2060;

RX MEDLINE=2056686; PubMed=1114919;

RA Sy A., Giraud E., Jourand P., Garcia N., Willems A., de lajudie P.,

RT "Methylobacterium nodulans" bacteria nodulate and fix nitrogen in

RT symbiosis with legumes.;

RL J. Bacteriol. 183:214-220(2001).

DR EMBL; AF220764; AAC49450.1; --

DR HSSP; P38539; 4AAH.

DR InterPro: IPR001479; Bac_PQQ.

DR InterPro: IPR002372; Bac_PQQ_repeat.

DR Pfam; PF01011; Bacterial_PQQ_7.

DR PROSITE; PS00363; BACTERIAL_PQQ_1; 1.

DR PROSITE; PS00364; BACTERIAL_PQQ_2; 1.

SQ SEQUENCE 629 AA; 68798 MW; 3DB559D8373BB0B2 CRC64;

Query Match 20.7%; Score 103; DB 2; Length 629;
Best Local Similarity 32.3%; Pred. No. 0.0087; Indels 8; Gaps 4;
Matches 30; Conservative 18; Mismatches 37;

OY 4 TSLMASAGALALAAPAFACVPTVDL--ANPPAGEMISTGONCENTRHSPLQITTE 62
DB 11 TGVSAAALALAPLPLGFS-----ALANDKVLKSKSDGVMVPGKNYDSDNSKILQINAE 66
OY 63 NVGQLOLVW--ARGMOPGKVOVTPILHDGVMY 93
DB 67 NVKNLKVSWQSTGLNGH-EGAPLVVDGTMV 98

RESULT 13

O9EYW8 PRELIMINARY; PRT; 601 AA.

AC O9EYW8: 01-MAR-2001 (TREMblrel. 16, Created)
DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)
DE METANOL DEHYDROGENASE LARGE SUBUNIT-LIKE PROTEIN.
GN MXAF.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Sinorhizobium.
NCBI_TaxID=382;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RM1021;
RA Fennel B.J., Tiwari R.P., Dilworth M.J.;
RT "Regulation of C1 assimilation in Sinorhizobium meliloti.";
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF309488; AAC31643.1; --
DR HSSP; P38539; 4AAH.
DR InterPro: IPR002372; Bac_PQQ_repeat.
DR Pfam; PF01011; Bacterial_PQQ_7.
SQ SEQUENCE 601 AA; 65686 MW; 09F743726B2E2A83 CRC64;

Query Match 19.9%; Score 99; DB 2; Length 601;
Best Local Similarity 34.5%; Pred. No. 0.022; Indels 6; Gaps 4;
Matches 29; Conservative 17; Mismatches 32;

OY 14 LALAAPAFACVPTVDL--LAMPAGEMISTGONCENTRHSPLQITTEENGQOLVW 71
DB 8 LAIMSTGGAGVAFANDELKIDLD--NOMALOTGDYANLRSKIDQINKDNVGLQVAM 66
OY 72 --ARGMOPGKVOVTPILHDGVMY 93
DB 67 TFSTGVLRGH-EGSPLVIGDLMY 89

RESULT 14
ID 092WY9 PRELIMINARY; PRT: 601 AA.
AC 092WY9;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE PUTATIVE METHANOL DEHYDROGENASE PROTEIN, LARGE SUBUNIT (EC 1.1.99.8).
GN SMO20173.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Sinorhizobium.
OX NCBI_TaxID=382;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1021;
RX MEDLINE=21396508; PubMed=11481431;
RA Flan T.M., Weldner S., Wong K., Buhrmester J., Chain P.,
RA Vorholter F.J., Hernandez-Lucas I., Becker A., Cowie A., Gouzy J.,
RA Golding B., Puehler A.;
RT "The complete sequence of the 1.683-kb pSYM megaplasmid from the N2-
fixing endosymbiont Sinorhizobium meliloti.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:9889-9894(2001).
DR EMBL: AL603642; CAC48573.1; -
KW Oxidoreductase; Plasmid; Hypothetical protein; Complete proteome.
SQ SEQUENCE 601 AA; 65759 MW; D73424FEED15ADBE CRC64;

Query Match 19.9%; Score 99; DB 16; Length 601;
Best Local Similarity 34.3%; Pred. No. 0.022;
Matches 29; Conservative 17; Mismatches 32; Indels 6; Gaps 4;

QY 14 LALLAFAFAQVPTDEL--LANPAGENSISGONENYRHSPLQITTEHYGOLQVW 71
Db 8 LAIMSIGGAGVAFANDELKLDLP--NQMAIQGTANLRLSKLDQINKDNKGLQVAV 66
QY 72 --ARGMOPGVQVPTPLIHDCVWYL 93
Db 67 TETGVLRGH-EGSPVIGDLMYV 89

RESULT 15
ID 024759 PRELIMINARY; PRT: 633 AA.
AC 024759;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE METHANOL DEHYDROGENASE ALPHA SUBUNIT PRECURSOR (EC 1.1.99.8).
GN KXAF.
OS Hyphomicrobium methylotrophum.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Hyphomicrobium group; Hyphomicrobium.
OX NCBI_TaxID=84;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GM2;
RX MEDLINE=97457202; PubMed=9311140;
RA Tanaka Y., Yoshida T., Watanabe K., Izumi Y., Mitsunaga T.;
RT "Cloning and analysis of methanol oxidation genes in the methylotroph
Hyphomicrobium methylotrophum GM2.";
RL FEMS Microbiol. Lett. 154:397-401(1997).
DR EMBL: AB004097; BAA23272.1; -
DR HSSP: P38539; 4AAH.
DR InterPro: IPR001479; Bac_P00.
DR InterPro: IPR002372; Bac_P00_repeat.
DR Pfam: PF01011; Bacterial_P00_7.
DR PROSITE: PS00363; BACTERIAL_P00_1.
DR PROSITE: PS00364; BACTERIAL_P00_2; 1.

KW Signal: Oxidoreductase.
FT SIGNAL 1 34 POTENTIAL.
FT CHAIN 35 633 POTENTIAL.
SQ SEQUENCE 633 AA; 69853 MW; B47A23A2779E2C1B CRC64;

Query Match 18.5%; Score 92; DB 2; Length 633;
Best Local Similarity 28.3%; Pred. No. 0.14;
Matches 26; Conservative 20; Mismatches 40; Indels 6; Gaps 3;

QY 6 LMASAGALALAAPFAQVPTDELANPAG--WISYGNENYRHSPLQITTEHY 64
Db 12 LMSASCAIIVALQVYASSAYANDKLIELSKNENWVPGKRYDSNNYSESTQVAAENV 71
QY 65 GQQLVYARGMOPGV---QVPTPLIHDCVWYL 93
Db 72 KQLHAMS--FTGELHGHGAPLVIGDMYV 101

Search completed: May 24, 2002, 10:17:28
Job time: 741 sec

5
r Fri May 24 11:27:24 2002

us-08-934-506a-5_copy_1_95.rspt

Page 7

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2002, 10:04:59 ; Search time 117.28 Seconds
(without alignments)
118.385 Million cell updates/sec

Title: US-08-934-506a-5_COPY_1_125

Perfect score: 657
Sequence: 1 MKPTSLMASGALALLAAP.....KTGDLWEHRRLPNTATLN 125

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: A-Geneseq.032802.*
2: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1980.DAT.*
3: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1981.DAT.*
4: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1982.DAT.*
5: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1983.DAT.*
6: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1984.DAT.*
7: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1985.DAT.*
8: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1986.DAT.*
9: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1987.DAT.*
10: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1988.DAT.*
11: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1989.DAT.*
12: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1990.DAT.*
13: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1991.DAT.*
14: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1992.DAT.*
15: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1993.DAT.*
16: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1994.DAT.*
17: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1995.DAT.*
18: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1996.DAT.*
19: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1997.DAT.*
20: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1998.DAT.*
21: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1999.DAT.*
22: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA2000.DAT.*
23: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA2001.DAT.*
24: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	657	100.0	579	19	AAW37873 Alcohol and/or ald
2	587	89.3	579	19	AAW37874 Alcohol and/or ald
3	548	83.4	579	19	AAW37876 Alcohol and/or ald
4	534.5	81.4	578	19	AAW37875 Alcohol and/or ald
5	189.5	28.8	754	21	AAW35987 Sorbitol dehydrog
6	183	27.9	738	12	AAW35993 A.altocegenes me
7	177	26.9	738	13	AAW35992 ADH complex protei
8	150.5	22.9	742	11	AAW35995 Amino acid sequenc
9	145.5	22.1	740	20	AAW35996 Novel human diagen
10	129	19.6	948	22	AAW35999 Sorbitol dehydrog
11					Novel human diagen

12	129	19.6	1510	22	ABG21573
13	76	11.6	824	18	AAW09614
14	73.5	11.2	1229	22	ABW71736
15	72.5	11.0	343	19	AAW35001
16	72.5	11.0	462	17	AAW05395
17	71	10.8	526	22	AAW99895
18	70	10.7	1016	22	AAW34349
19	69.5	10.6	275	21	AAW42781
20	69.5	10.6	1419	22	ABW59208
21	69	10.5	711	22	ABW69743
22	69	10.5	957	22	ABW62401
23	69	10.5	1029	22	ABW6194
24	69	10.5	1029	22	ABW628658
25	68.5	10.4	1029	22	ABW61927
26	68	10.4	424	22	ABW59716
27	67.5	10.3	171	22	AAW32330
28	67.5	10.3	275	21	AAW16656
29	67.5	10.3	1041	22	ABW31594
30	67.5	10.3	1224	18	AAW6735
31	67	10.2	1224	18	AAW63034
32	67	10.2	159	19	AAW4325
33	67	10.2	974	19	AAW49079
34	67	10.2	1407	22	ABW28517
35	67	10.2	1429	22	ABW58779
36	66.5	10.1	333	22	ABW63676
37	66.5	10.1	384	22	ABW65251
38	66.5	10.1	551	22	ABW85252
39	66.5	10.1	551	22	ABW85252
40	66.5	10.1	577	22	ABW08719
41	66	10.0	227	22	AAW56221
42	66	10.0	620	21	AAW77208
43	66	10.0	620	21	ABW6214
44	66	10.0	673	22	ABW59674
45	65.5	10.0	271	18	AAW28260

ALIGNMENTS

RESULT	1
AAW37873	AAW37873 standard; Protein; 579 AA.
XX	XX
AC	AAW37873;
XX	XX
DT	10-AUG-1998 (first entry)
DE	Alcohol and/or aldehyde dehydrogenase A amino acid sequence.
XX	XX
KW	Alcohol/aldehyde dehydrogenase A enzyme; recombinant organism;
KW	aldehyde; ketone; carboxylic acid; L-sorbose; D-sorbitol;
KW	2-keto-L-gulonate acid; L-ascorbic; inhibition.
XX	XX
OS	Gluconobacter oxydans.
XX	XX
FT	Key
FT	Peptide
FT	Protein
FT	Location/Qualifiers
FT	1..23
FT	/note="signal peptide"
FT	24..579
FT	/note="mature protein"
XX	XX
PN	EP832974-A2.
XX	XX
PD	01-APR-1998.
XX	XX
PF	11-SEP-1997; 97EP-0115801.
XX	XX
PR	19-SEP-1996; 96EP-0115001.
XX	XX
PA	(HOFF) HOFFMANN LA ROCHE & CO AG F.
XX	XX
PI	Asakura A, Hoshino T, Ojima S, Shinjoh M, Tomiyama N;
XX	XX

Novel human diagen
Pathogenic Staphyl
Drosophila melanog
Bacillus thiermoide
Human SH3P1 prote
Xanthomonas sp. ch
Staphylococcus aur
Arabidopsis thailia
Drosophila melanog
Drosophila melanog
Drosophila melanog
Novel human diagen
Novel human diagen
Acromonium cellulo
Drosophila melanog
Novel human secret
Arabidopsis thailia
Novel human diagen
Staphylococcus car
Mycobacterium tube
Mycobacterium tube
Solium tuberosum
Novel human diagen
Drosophila melanog
Drosophila melanog
Plant porobilinog
Plant thioedoxin-
Novel human diagen
Propionibacterium
S. venezuelae deso
Drosophila melanog
Drosophila melanog
Amino acid sequenc

DR WPI: 1998-195228/18.
 DR N-PSDB: AAV29051.
 XX
 PT Recombinant Gluconobacter oxydans alcohol and/or aldehyde
 PT dehydrogenase enzyme(s) - useful for converting substrate(s), e.g.
 PT L-sorbose or D-sorbitol to 2-keto-L-gulononic acid
 XX
 PS Claim 1; Pages 35-37; 59pp; English.
 XX
 CC This is the amino acid sequence for the Gluconobacter oxydans
 CC alcohol and/or aldehyde dehydrogenase A' enzyme. The enzymes
 CC or recombinant organisms can be used to convert suitable substrates
 CC to aldehydes, ketones or carboxylic acids, especially to convert
 CC L-sorbose or D-sorbitol to 2-keto-L-gulononic acid, which can be
 CC converted to L-ascorbic acid by standard procedures. The derivatives
 CC of ADH enzymes have desired substrate specificity, higher affinity
 CC to a substrate, lower affinity to an inhibitory compound, higher
 CC stability against temperature and/or pH and higher catalytic speed.
 CC
 SQ Sequence 579 AA;

Query Match 100.0%; Score 657; DB 19; Length 579;
 Best Local Similarity 100.0%; Pred. No. 2,2e-64;
 Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MKPTSLIMASAGALIALAAPAFQVTPVDELLANPPAGEMISYCONENYRHSPLTQIT 60
 Db 1 MKPTSLIMASAGALIALAAPAFQVTPVDELLANPPAGEMISYCONENYRHSPLTQIT 60
 QY 61 TENVGQLQVWARGMOPGKQVTPPLIHGVMYLANPGDVIAQADKTDGLIWEHRRLPN 120
 Db 61 TENVGQLQVWARGMOPGKQVTPPLIHGVMYLANPGDVIAQADKTDGLIWEHRRLPN 120
 QY 121 IATLN 125
 Db 121 IATLN 125

RESULT 2
 AAW37874
 ID AAW37874 standard; Protein: 579 AA.
 AC AAW37874;
 XX
 DT 10-AUG-1998 (first entry)
 DE
 XX Alcohol and/or aldehyde dehydrogenase A' amino acid sequence.
 KW Alcohol/aldehyde dehydrogenase A' enzyme; recombinant organism;
 KW aldehyde; ketone; carboxylic acid; L-sorbose; D-sorbitol;
 KW 2-keto-L-gulononic acid; L-ascorbic; inhibition.
 XX
 OS Gluconobacter oxydans.
 XX
 FH Key
 FT Peptide 1..23 Location/Qualifiers
 FT /note= "signal peptide"
 FT Protein 24..579
 FT /note= "mature protein"
 XX
 PN EP832974-A2.
 PD 01-APR-1998.
 PF 11-SEP-1997; 97EP-0115801.
 XX
 PR 19-SEP-1996; 96EP-0115001.
 XX
 PA (HOFF) HOFFMANN LA ROCHE & CO AG F.
 PI Asakura A, Hoshino T, Ojima S, Shinjoh M, Tomiyama N;
 XX

DR WPI: 1998-195228/18.
 DR N-PSDB: AAV29052.
 XX
 PT Recombinant Gluconobacter oxydans alcohol and/or aldehyde
 PT dehydrogenase enzyme(s) - useful for converting substrate(s), e.g.
 PT L-sorbose or D-sorbitol to 2-keto-L-gulononic acid
 XX
 PS Claim 1; Pages 38-40; 59pp; English.
 XX
 CC This is the amino acid sequence for the Gluconobacter oxydans
 CC alcohol and/or aldehyde dehydrogenase A' enzyme. The enzymes
 CC or recombinant organisms can be used to convert suitable substrates
 CC to aldehydes, ketones or carboxylic acids, especially to convert
 CC L-sorbose or D-sorbitol to 2-keto-L-gulononic acid, which can be
 CC converted to L-ascorbic acid by standard procedures. The derivatives
 CC of ADH enzymes have desired substrate specificity, higher affinity
 CC to a substrate, lower affinity to an inhibitory compound, higher
 CC stability against temperature and/or pH and higher catalytic speed.
 CC
 SQ Sequence 579 AA;

Query Match 89.3%; Score 587; DB 19; Length 579;
 Best Local Similarity 91.2%; Pred. No. 1.3e-56;
 Matches 114; Conservative 3; Mismatches 8; Indels 0; Gaps 0;
 QY 1 MKPTSLIMASAGALIALAAPAFQVTPVDELLANPPAGEMISYCONENYRHSPLTQIT 60
 Db 1 MKPTSLIMASAGALIALAAPAFQVTPVDELLANPPAGEMISYCONENYRHSPLTQIT 60
 QY 61 TENVGQLQVWARGMOPGKQVTPPLIHGVMYLANPGDVIAQADKTDGLIWEHRRLPN 120
 Db 61 TENVGQLQVWARGMOPGKQVTPPLIHGVMYLANPGDVIAQADKTDGLIWEHRRLPN 120
 QY 121 IATLN 125
 Db 121 IATLN 125

RESULT 3
 AAW37876
 ID AAW37876 standard; Protein: 579 AA.
 AC AAW37876;
 XX
 DT 10-AUG-1998 (first entry)
 DE
 XX Alcohol and/or aldehyde dehydrogenase B amino acid sequence.
 KW Alcohol/aldehyde dehydrogenase B enzyme; recombinant organism;
 KW aldehyde; ketone; carboxylic acid; L-sorbose; D-sorbitol;
 KW 2-keto-L-gulononic acid; L-ascorbic; inhibition.
 XX
 OS Gluconobacter oxydans.
 XX
 FH Key
 FT Peptide 1..23 Location/Qualifiers
 FT /note= "signal peptide"
 FT Protein 24..579
 FT /note= "mature protein"
 XX
 PN EP832974-A2.
 PD 01-APR-1998.
 PF 11-SEP-1997; 97EP-0115801.
 XX
 PR 19-SEP-1996; 96EP-0115001.
 XX
 PA (HOFF) HOFFMANN LA ROCHE & CO AG F.
 PI Asakura A, Hoshino T, Ojima S, Shinjoh M, Tomiyama N;
 XX

DR MPI: 1998-195228/18.
DR N-PSDS: AAV29054.
XX
XX Recombinant Gluconobacter oxydans alcohol and/or aldehyde
XX dehydrogenase enzyme(s) - useful for converting substrate(s), e.g.
PT L-sorbose or D-sorbitol to 2-keto-L-gulononic acid
XX
XX Claim 1; Pages 44-46; 59pp; English.
PS
XX
XX This is the amino acid sequence for the gluconobacter oxydans
CC alcohol and/or aldehyde dehydrogenase B enzyme. The enzymes
CC or recombinant organisms can be used to convert suitable substrates
CC to aldehydes, ketones or carboxylic acids, especially to convert
CC L-sorbose or D-sorbitol to 2-keto-L-gulononic acid, which can be
CC converted to D-xescorbic acid by standard procedures. The derivatives
CC of ADH enzymes have desired substrate specificity, higher affinity
CC to a substrate, lower affinity to an inhibitory compound, higher
CC stability against temperature and/or pH and higher catalytic speed.
XX
XX Sequence 579 AA;
SQ

Query Match	83.4%;	Score 548;	DB 19;	Length 579;
Best Local Similarity	81.6%;	Fred. No. 2.9e-52;		
Matches 102;	Conservative 10;	Mismatches 13;	Indels 0;	Gaps 0;
Qy	1 MKPSLILMASAGALATLLAPAFACQVTPVDELLAPPAGEIMISYGCNOENYRHSPLTQIT 60			
Db	1 mmpLLILrLrSaavllltpaaafagvtrpLldellamppegewlgrngeryrnsplltqit 60			
Qy	61 TENYQGLQQLVMWAGKMPGKVQVPTLIDGCVMTLANPGGVIOAIDAKTGDILWEHRRQLPN 120			
Db	61 adnvgqqlqvlwargmeagavqvcymihdvmvylangpvglaqladqtrgdllwehrrqlpa 120			
Qy	121 TATLNL 125			
Db	121 valn 125			

RESULT	4
AAW37875	
ID	AAW37875 standard; Protein; 578 AA.
AC	
XX	AAW37875;
XX	
DT	10-AUG-1998 (first entry)
XX	
DE	Alcohol and/or aldehyde dehydrogenase A'' amino acid sequence.
XX	
KW	Alcohol/aldehyde dehydrogenase A'' enzyme; recombinant organism
KW	aldehyde; ketone; carboxylic acid; L-sorbose; D-sorbitol;
KW	2-keto-L-gulonate acid; L-ascorbic; inhibition.
XX	
OS	Gluconobacter oxydans.
XX	
FH	Key
FT	Peptide
FT	1..23
FT	/note= "signal peptide"
FT	24..578
FT	/note= "mature protein"
XX	
PN	EP632974-A2.
XX	
PD	01-APR-1998.
XX	
PE	11-SEP-1997; 97EP-0115801.
XX	
PR	19-SEP-1996; 96EP-0115001.
XX	
PA	(HOFF) HOFFMANN LA ROCHE & CO AG F.
XX	
PI	Asakura A, Hoshino T, Ojima S, Shinjoh M, Tomiyama N;
XX	

DR	WPI; 1998-195228/18.
DR	N-PSDB; AAV29053.
XX	Recombinant Gluconobacter oxydans alcohol and/or aldehyde
PT	dehydrogenase enzyme(s) - useful for converting substrate(s), e.g.
PT	L-sorbose or D-sorbitol to 2-keto-L-gulononic acid
XX	
XX	Claim 1; Pages 41-43; 59pp; English.
XX	
CC	This is the amino acid sequence for the Gluconobacter oxydans
CC	alcohol and/or aldehyde dehydrogenase A'' enzyme. The enzymes
CC	or recombinant organisms can be used to convert suitable substrates
CC	to aldehydes, ketones or carboxylic acids, especially to convert
CC	L-sorbose or D-sorbitol to 2-keto-L-gulononic acid, which can be
CC	converted to L-ascorbic acid by standard procedures. The derivativ
CC	of AAOH enzymes have desired substrate specificity, higher affinity
CC	to a substrate, lower affinity to an inhibitory compound, higher
CC	stability against temperature and/or pH and higher catalytic speed.
XX	
SQ	Sequence 578 AA;

[illegible]

PT	Novel membrane-bound sorbitol dehydrogenase nucleic acid molecule isolated from Gluconobacter suboxydans useful for the fermentative p
XX	
DR	N-PSDB; AAC83153.
XX	
PI	Choi E, Rhee S, Lee E;
XX	
PA	(CHOI/) Choi E,
PA	(RHEE/) Rhee S,
PA	(LEE/) Lee E.
XX	
PR	23-APR-1999; 99WO-IB00736.
XX	
PF	23-APR-1999; 99WO-IB00736.
XX	
PD	02-NOV-2000.
XX	
PN	WO200065066-A1.
XX	
OS	Gluconobacter oxydans.
KW	L-sorbose production; 2-keto-L-gulononic acid.
KW	Sorbitol dehydrogenase; SDH; Gluconobacter oxydans; D-sorbitol;
DE	Sorbitol dehydrogenase subunit 1 amino acid sequence.
XX	
DI	01-MAR-2001 (first entry)
AC	AAB35987;
XX	
ID	AAB35987 standard; Protein; 754 AA.
RESULT	5
AAB35987	

PF Production of 2-keto-L-gulononic acid and L-sorbose from D-sorbitol
 XX
 PS Claim 1; Fig 8; 96pp; English.
 CC This invention relates to an isolated membrane-bound sorbitol
 CC dehydrogenase (SDH) from Gluconobacter oxydans. The invention includes
 CC polynucleotides (AAC83153 - AAC83154) encoding the 3 subunits of SDH
 CC (AAB35987 - AAB35989). Also included in the invention are two
 CC polynucleotide sequences AAC83156 and AAC83157 which contain the subunit
 CC coding regions. The sorbitol dehydrogenase polynucleotide sequences are
 CC useful for producing L-sorbose from D-sorbitol and for increasing the
 CC production of 2-keto-L-gulononic acid by transforming a host cell,
 CC especially Gluconobacter with the DNA and selecting the transformed host
 CC cell. Note: An additional 66 DNA sequences (fragments of the SDH subunit
 CC genes) encoding fragments of SDH are specifically claimed, however these
 CC sequences are not given separately in the specification but are included
 CC in sequences AAC83156 and AAC83157.
 CC
 XX Sequence 754 AA:

Query Match 28.8%; Score 189.5; DB 21; Length 754;
 Best Local Similarity 39.0%; Pred. No. 3,1e-12;
 Matches 46; Conservative 17; Mismatches 46; Indels 9; Gaps 3;

OY 14 LALLAARAFQVTPV-----TDELLANPPAGEMISYGONENYRHSPLTQITTEYNGQ 66
 Db 17 lgcacaaatcatspvalaedtgaltadqhpqdwmsygrtysqrgpldtktdnasn 76
 OY 67 LQLVWARGMQPKQVQV-TPLIHGVMYLANPGDVIOAIDAKTGLIMEHRQLP-NIA 122
 Db 77 lklawhyldldtrnggegrplivdygvyattmwmkavdaatgkllwsydrvpqna 134

RESULT 6

AAR13993
 ID AAR13993 standard; Protein; 738 AA.

AC AAR13993;

DT 09-DEC-1991 (first entry)

DE A. altocetigenes membrane-bound ADH 72kd sub-unit.

KW alcohol dehydrogenase complex; carboxylic acid production.

OS Acetobacter altocetigenes.

PN EP448969-A.

PD 02-OCT-1991.

PF 26-FEB-1991; 91EP-0102793.

PR 26-MAR-1990; 90JP-0073440.

PR 26-FEB-1990; 90JP-0042391.

PA (NAKA-) NAKANO VINEGAR KK.

PI Tamaki T, Takemura H, Tayama K, Fukaya M, Okumura H;

DR WPI; 1991-289462/40.

DR N-PSDB; AAQ13580.

PT Gene for membrane-bound alcohol dehydrogenase complex - obtd.

PS from Acetobacter altocetigenes, used for prodn. of enzyme for

CC converting alcohol to acid
 CC Disclosure; Fig 3; 36pp; English.
 CC Total DNA was prepared from A. altocetigenes MH-24, digested with
 CC PstI and SmaI and ligated to PstI-SmaI cleaved pUC18. The ligation

CC mixture was used to transform E.coli JM109. Probes were designed
 CC based on the N-terminal amino acid sequence of the ADH complex
 CC isolated from A. altocetigenes (see AAQ13582-013584). The gene
 CC encoding the 72kd ADH subunit was isolated and sequenced. The
 CC directly sequenced N-terminal region of the purified 72kd sub-unit
 CC corresponds to the sequence beginning at residue 36 of the deduced
 CC sequence. This suggests that the first 35 N-terminal amino acids
 CC form a leader peptide involved in secretion of the mature 72kd
 CC protein. The deduced amino acid sequence has 77 per cent homology
 CC with the same enzyme from A. aceti K6033. See also AAQ13581.
 CC
 XX Sequence 738 AA;

Query Match 27.9%; Score 183; DB 12; Length 738;
 Best Local Similarity 33.3%; Pred. No. 1,6e-11;
 Matches 41; Conservative 23; Mismatches 49; Indels 10; Gaps 3;

OY 10 SAGALALLAARAFQVTPVDE-----LLANPPAGEMISYGONENYRHSPLTQITTT 61
 Db 17 tagtcaalltsqyaltmasaddqggaatgaalhadddpnmvtygrtysdgryspldqlnr 76
 OY 62 ENVCQLQVWARGMQPKQVQV-TPLIHGVMYLANPGDVIOAIDAKTGLIMEHRQLP- 119
 Db 77 snvgnlklawyldldtrnggegrplivdygvyattmwmkavdaatgkllwsydrvpq 136
 OY 120 NIA 122
 Db 137 nla 139

RESULT 7

AAR20192
 ID AAR20192 standard; Protein; 738 AA.

AC AAR20192;

DT 16-APR-1992 (first entry)

DE ADH complex protein (mol.wt. 72.000).

KW Alcohol dehydrogenase; acetic acid; fermentation.

OS Acetobacter altocetigenes NH-24.

PN JP03266988-A.

PD 27-NOV-1991.

PF 26-MAR-1990; 90JP-0073440.

PR 26-FEB-1990; 90JP-0042301.

PR 26-MAR-1990; 90JP-0073440.

PA (NAKA-) NAKANO SUTEN KK.

PI WPI; 1992-019325/03.

DR N-PSDB; AAQ20383.

PT Alcohol dehydrogenase complex structural gene - used in plasmid

PS and enhancing efficiency of acetic acid fermentation for

CC transformed acetic acid bacteria
 CC Disclosure; Fig 3(1-3); 21pp; Japanese.
 CC Acetobacter transformed with the sequence encoding this protein can
 CC enhance the efficiency of acetic acid fermentation. The ADH complex
 CC can be easily extracted from the bacteria and purified and it can be
 CC used for the determination of an alcohol.
 CC See also AAQ20383-84, and -86-88.

CC Sequence 738 AA;

Db 168 ltwagfndpqeIngfLlsadatpaea-Isprvadq-----dwpaygrngesqrfspkq 218
 QY 59 ITTEVNGQLQIYWA-----RGWOPGKY--QVTPLIHDGVMTLANPQDVIQAIIDAKTGD 109
 Db 219 InadvhnlkwaWvfrtgdkvkgndpgeInevtpikvgdLlyctahgrlfaldaaagk 278
 QY 110 LIWEHRROL 118
 Db 279 ekWnydpel 287

RESULT 10

AAW95019
 ID AAW95019 standard; Protein: 740 AA.

AC AAW95019;

DT 21-MAY-1999 (first entry)

DE Sorbitol dehydrogenase (SLDH) protein sequence.

KW Sorbitol dehydrogenase; SLDH; open reading frame; ORF2; L-sorbose;
 KM D-sorbitol; vitamin C; enzyme.

OS Gluconobacter suboxydans.

FX Key Location/Qualifiers

FT Peptide 1..24

FT /note= "signal sequence"

FT Protein 25..740

FT /note= "mature protein"

PN EP897984-A2.

PD 24-FEB-1999.

XX 13-AUG-1998; 98EP-0115231.

XX 21-AUG-1997; 97EP-0114432.

XX (HOFF) HOFFMANN LA ROCHE & CO AG F.

XX Hosino T, Miyazaki T, Ojima S, Shinjoh M, Tomiyama N;

DR WPI: 1999-134646/12.

DR N-PSDB: AAX21501.

XX New D-sorbitol dehydrogenase gene and recombinant protein - useful
 PT for production of L-sorbose, an intermediate in vitamin C production

XX Claim 1; Fig 3A-D; 39pp; English.

CC This represents a sorbitol dehydrogenase (SLDH) protein. The DNA
 CC encoding the SLDH enzyme also encodes an open reading frame (ORF2)
 CC product upstream of the SLDH open reading frame, needed for SLDH activity
 CC in vivo. Host cells transfected by a vector comprising the SLDH DNA
 CC sequence are used for the recombinant expression of the sorbitol
 CC dehydrogenase. Recombinant D-sorbitol dehydrogenase is useful for
 CC producing L-sorbose from D-sorbitol. L-sorbose is an important
 CC intermediate in vitamin C production.

CC Sequence 740 AA;

Query Match 22.1%; Score 145.5; DB 20; Length 740;
 Best Local Similarity 34.0%; Pred. No. 2.4e-07;
 Matches 34; Conservative 25; Mismatches 30; Indels 11; Gaps 5;

QY 23 AGVTVTDELLANPAGIEWISGQENYRHSPLTQITTEVNGQLQI--VWARGM--OPG 78

Db 86 sgvpmamqgsanpargdvwgrygriddhqttrypslseipenaskrkvafyhngsyprrp 145

QY 79 KV-----QVTPLI-HHDGVMTLANPQDVIQAIIDAKTGLIWI 112

Db 146 qvkwaaettpikvgdgllycsamndlk-ldpatgkqLw 184

RESULT 11

ABG25909
 ID ABG25909 standard; Protein: 948 AA.

AC ABG25909;

DT 18-FEB-2002 (first entry)

DE Novel human diagnostic protein #25900.

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;

KW food supplement; medical imaging; diagnostic; genetic disorder.

OS Homo sapiens.

PN WO200175067-A2.

PD 11-OCT-2001.

PF 30-MAR-2001; 2001WO-US08631.

PR 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

PA (HYSE-) HYSEQ INC.

PI Dmanac RT, Liu C, Tang YF;

DR WPI: 2001-639362/73.

DR N-PSDB: AAS90096.

XX New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity

PS Claim 20; SEQ ID No 56268; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pat_sequences.

CC Sequence 948 AA;

Query Match 19.6%; Score 129; DB 22; Length 948;

Best Local Similarity 29.9%; Pred. No. 2.2e-05;
 Matches 43; Conservative 17; Mismatches 30; Indels 54; Gaps 8;

QY 6 LMAAS-----AGALALAAPFAOVTPVTDELLANPAGIEWISGQENYRHSPLTQ 58

OY 71 WARGMGKQVVTPLJHD-----GVMTLNP--GDVIO-AIDAKT----- 107
 DB 733 -----eyptcttprhtrdtklptgekeevpdkpiknpecdgavrvpvdsvtkygvv 787
 OY 108 -GDILWEHRRLP 119
 DB 788 kgdsive-keelp 799

RESULT 14

ABR71736
 ID ABR71736 standard; Protein; 1229 AA.

AC ABR71736;

DT 26-MAR-2002 (first entry)

DE Drosophila melanogaster polypeptide SEQ ID NO 42000.

KW Drosophila; developmental biology; cell signalling; insecticide; pharmaceutical.

OS Drosophila melanogaster.

PN WO200171042-A2.

PD 27-SEP-2001.

PF 23-MAR-2001; 2001WO-US09231.

PR 23-MAR-2000; 2000US-191637P.

PR 11-JUL-2000; 2000US-0614150.

PA (PEKE) PE CORP NY.

PI Venter JC, Adams M, Li PWD, Myers EW;

DR WPI; 2001-656860/75.

DR N-PSDB; ABL15839.

PT New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signalling and cell-cell interactions -

PS Disclosure; SEQ ID NO 42000; 21pp + Sequence Listing; English.

CC The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA sequences (ABL01840-ABL16175) and the encoded proteins

CC (ABR71737-ABR72072).
 CC The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

SO Sequence 1229 AA;

Query Match 11.2%; Score 73.5; DB 22; Length 1229;
 Best Local Similarity 31.2%; Pred. No. 47;
 Matches 25; Conservative 13; Mismatches 33; Indels 9; Gaps 4;

OY 8 WASGALALAAPFAOV-TPTVDLILNP--PAGEWISYGOQEN--YRHSPLTQITTE 62
 DB 958 yvskpavaisaapaiaakvavptvshstgptlgyvgvygssghgvgllapltltsa 1017

OY 63 NV-----GOLQLVWARGMQPG 78

DB 1018 payalggkiasstcaylqag 1037

RESULT 15
 AAW35001
 ID AAW35001 standard; Protein; 343 AA.

AC AAW35001;

DT 21-MAY-1998 (first entry)

DE Bacillus thermooleovorans endoglucanase.

KW Endoglucanase; cellulase; carboxymethylcellulose; cellulose; biomass; beta-1,4-glycosidic bond; hydrolysis; saccharification; thermostable enzyme; thermophilic; glycosidase.

OS Bacillus thermooleovorans (Clone 68GCl).

PN WO9744361-A1.

PD 27-NOV-1997.

PF 22-MAY-1997; 97WO-US08793.

PR 22-MAY-1996; 96US-0651572.

PA (RECO-) RECOMBINANT BIOCATALYSIS INC.

PI Lam DE, Mathur EJ;

DR WPI; 1998-018435/02.

DR N-PSDB; AAT94209.

PT Endoglucanase(s), preferably from archaeal bacterium, AEPIT 1a - useful to degrade carboxymethylcellulose and hydrolyse of beta-1,4-glycosidic bonds in cellulose

PS Claim 1; Fig 10; 164pp; English.

CC This protein comprises an endoglucanase of Bacillus thermooleovorans (Clone 68GCl) that is capable of degrading carboxymethylcellulose and of hydrolysing the beta-1,4-glycosidic bonds in cellulose. It has homology to an endoglucanase of archaeobacterium AEPIT1a (see AAW34985). It can be produced from native cells or from recombinant host cells, especially prokaryotic host cells transformed with a plasmid or virus-derived vector including the endoglucanase DNA (see AAT94209). 24 Endoglucanases (see AAW34986-W35008) are claimed. They can be used to degrade cellulose for the conversion of plant biomass into fuels and chemicals, for use in detergents, textiles, animal feed, waste treatment, and in the fruit juice and brewing industries for the clarification and extraction of juices.

SO Sequence 343 AA;

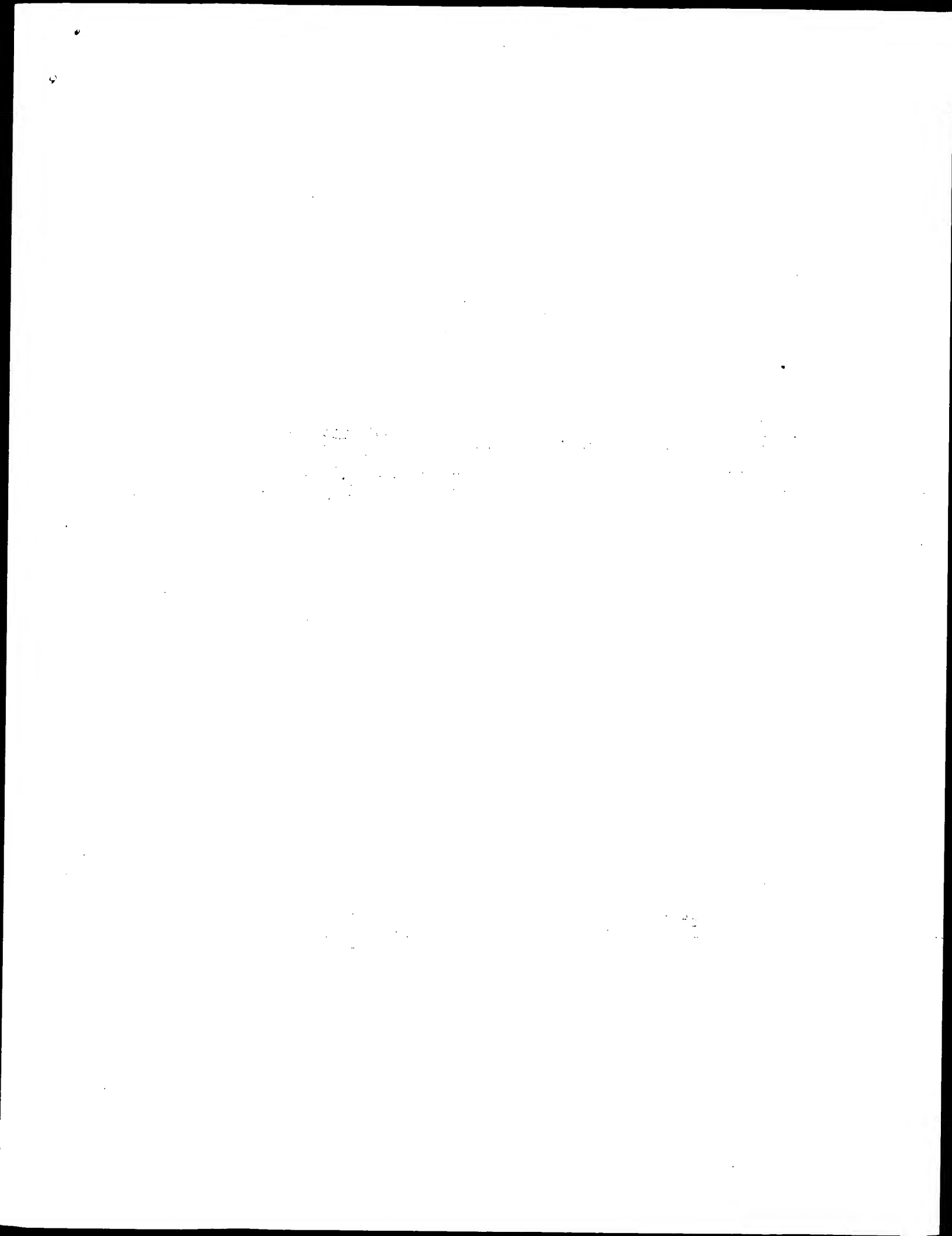
Query Match 11.0%; Score 72.5; DB 19; Length 343;
 Best Local Similarity 35.3%; Pred. No. 11;
 Matches 18; Conservative 6; Mismatches 26; Indels 1; Gaps 1;

OY 28 VTDELLANPPAGEWISYGOQENYRHSPLTQITTEVGOLOLVWARGMQPG 78
 DB 106 vlenasamplgegwekqvtkwesfalddattfkhkxyvvaq-kdpg 155

Search completed: May 24, 2002, 10:05:00
 Job time: 239 sec

Fri May 24 11:27:12 2002

us-08-934-506a-5_copy_1_125.rag




```

? NUMBER OF SEQUENCES: 4
? CORRESPONDENCE ADDRESS:
? STREET: 600 Third Avenue
? CITY: New York
? STATE: New York
? COUNTRY: USA
? ZIP: 10016-2088
? COMPUTER READABLE FORM:
? MEDIUM TYPE: 3.5 inch, 0.72 mb
? COMPUTER: IBM PC compatible (NEC PC-9801 ES)
? OPERATING SYSTEM: MS DOS
? SOFTWARE: ASCII Form
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/07/985,458
? FILING DATE: 19921203
? CLASSIFICATION: 435
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 07/658,221
? FILING DATE: 20-FEB-1991
? APPLICATION NUMBER: 73440/1990
? FILING DATE: 26-MAR-1990
? ATTORNEY/AGENT INFORMATION:
? NAME: Goodman, Herbert
? REGISTRATION NUMBER: 17081
? REFERENCE/DOCKET NUMBER: 910134/HG
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (212)972-1400
? TELEFAX: (212)370-1622
? TELEX: 236268
? INFORMATION FOR SEQ ID NO: 3:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 738 amino acids
? TYPE: amino acid
? TOPOLOGY: unknown
? MOLECULE TYPE: protein
? FEATURE:
? NAME/KEY: MATURE PEPTIDE
? LOCATION: 36 to 738
? IDENTIFICATION METHOD: N-terminal sequences of the
? IDENTIFICATION METHOD: Purified protein having a molecular weight of about
? IDENTIFICATION METHOD: 72,000
? ORIGINAL SOURCE:
? ORGANISM: Acetobacter alioacetigenes
? STRAIN: MH-24
? PUBLICATION INFORMATION:
? AUTHORS: Tanaka, Toshimi;
? AUTHORS: Fukaya, Masahiro;
? AUTHORS: Takemura, Hiroshi;
? AUTHORS: Tayama, Kenji;
? AUTHORS: Kawamura, Hajime;
? AUTHORS: Kawamura, Yoshio;
? AUTHORS: Nishiyama, Makoto;
? AUTHORS: Horiuchi, Suenaru and
? AUTHORS: Beppu, Teruhiko
? TITLE: Cloning and Sequencing of the Gene Cluster
? TITLE: Encoding Two Subunits of Membrane-Bound
? TITLE: Alcohol Dehydrogenase from Acetobacter
? JOURNAL: Biochimica et Biophysica Acta.
? VOLUME: 1088
? PAGES: 292-300
? DATE: 1991
? US-07-985-458-3

```

```

Query Match          27.9%; Score 183; DB 1; Length 738;
Best Local Similarity 33.3%; Pred. No. 6.1e-13;
Matches 41; Conservative 23; Mismatches 49; Indels 10; Gaps 3;
QY 10 SAGALALAAPFAQVTPPTDE-----LLANPPAGEMTSYGNQENYRHSPLTQTTT 61
DB 17 TAGTICALISGYATMASDGGATGATRAIIHADHPGNMWTGRTYSQDRYSPIDQINR 76

```

```

QY 62 ENVGQLQVWARGMQPGKQVQV-TPLIHDGVWYLANPQGVQIAIDAKTGDLWEHRRQLP-119
DB 77 SNVGKLAWYLDLDTNRNGQCGTFLYIDGVVYATTNMSMKAVDAATGKILMSYDPRVPG 136
QY 120 NIA 122
DB 137 NIA 139
RESULT 3
US-09-296-284-25
? Sequence 25, Application US/09296284A
? Patent No. 6204040
? GENERAL INFORMATION:
? APPLICANT: Choi, Eun-Sung
? APPLICANT: Rhee, Sang-Ki
? APPLICANT: Lee, Eun-Hae
? TITLE OF INVENTION: Glucobacter Suboxydans Sorbitol Dehydrogenase, Genes
? FILE REFERENCE: 1533.0870000
? CURRENT APPLICATION NUMBER: US/09/296,284A
? CURRENT FILING DATE: 1999-04-22
? NUMBER OF SEQ ID NOS: 87
? SOFTWARE: Patentin Ver. 2.0
? SEQ ID NO 25
? LENGTH: 720
? TYPE: PRT
? ORGANISM: Glucobacter suboxydans
US-09-296-284-25

```

```

Query Match          27.5%; Score 180.5; DB 4; Length 720;
Best Local Similarity 40.4%; Pred. No. 1.2e-12;
Matches 40; Conservative 18; Mismatches 36; Indels 5; Gaps 3;
QY 26 TPVTELLANPPAGEMTSYGNQENYRHSPLTQTTENVGQLQVWARGMQPGKQVQV-TP 84
DB 5 TAITN---ADQHPGDMWSGRTYSFORSPLDQITKDNASNLKLAHWYDLDTNRNGQESTP 61
QY 85 LHDGVWYLANPQGVQIAIDAKTGDLWEHRRQLP-NIA 122
DB 62 LIVDGVWYATTNMSMKALDAATGKILMSYDPKVGANIA 100

```

```

RESULT 4
US-09-136-251-2
? Sequence 2, Application US/09136251A
? Patent No. 6127156
? GENERAL INFORMATION:
? APPLICANT: HOSHINO, Tatsuo
? APPLICANT: MIYAZAKI, Taro
? APPLICANT: OJIMA, Setsuko
? APPLICANT: SHINJOH, Masako
? TITLE OF INVENTION: D-SORBITOL DEHYDROGENASE GENE
? FILE REFERENCE: D-Sorbitol Dehydrogenase Gene
? CURRENT APPLICATION NUMBER: US/09/136,251A
? CURRENT FILING DATE: 1998-08-19
? EARLIER APPLICATION NUMBER: EP 97114432.4
? NUMBER OF SEQ ID NOS: 9
? SOFTWARE: Patentin Ver. 2.1
? SEQ ID NO 2
? LENGTH: 739
? TYPE: PRT
? ORGANISM: Glucobacter suboxydans
? FEATURE:
? NAME/KEY: SIGNAL
? LOCATION: (1)..(24)
US-09-136-251-2

```


Query Match 10.2%; Score 67; DB 3; Length 159;
Best Local Similarity 21.8%; Pred. No. 2.4;
Matches 36; Conservative 18; Mismatches 49; Indels 62; Gaps 7;

QY 1 MKPISLMSAGALALAAFAVPTVDLLANPAGWISYGNQENYRHSPLTQIT 60
DB 6 MKTAVAVVAAAIATPAAPALAPITGKL-----GSELTM-- 43

QY 61 TENYGOLOLVW-----ARGMOPG-----KVQVTPLIH-----DG 89
DB 44 TDYGVGVLGKVKVSDKSTAVIGYPPVAGVWEATATVNAIRGSYTPAVSOFNATRADG 103

QY 90 VY-----LANP-----GDVQAIIDAKTGDLIWEHRRPLNATL 125
DB 104 INRYVLMQAGPDLTSCATIPQEGESTGKIXFDVYGPSPLVAMN 148

RESULT 10
US-08-868-786-6
; Sequence 6, Application US/08868786
; Patent No. 5998701

GENERAL INFORMATION:
APPLICANT: Kawchuk, Lawrence M.
APPLICANT: Armstrong, John
APPLICANT: Lynch, Dermot
APPLICANT: Knowles, Richard
TITLE OF INVENTION: Potatoes Having Improved Quality
TITLE OF INVENTION: Characteristics and Methods for Their Production
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
STREET: 5730 Manhattan Circle, Suite 201
CITY: Boulder
STATE: Colorado
COUNTRY: USA

ZIP: 80803
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/868,786
FILING DATE: 04-JUN-1997
CLASSIFICATION: 800

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/036,946
FILING DATE: 10-FEB-1997
ATTORNEY/AGENT INFORMATION:
NAME: Sullivan, Sally A
REGISTRATION NUMBER: 32,064

REFERENCE/DOCKET NUMBER: 8-97
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 499-8080
TELEFAX: (303) 499-8089
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 974 amino acids
TYPE: amino acid
TOPOLOGY: linear

MOLECULE TYPE: protein
US-08-868-786-6

Query Match 10.2%; Score 67; DB 2; Length 974;
Best Local Similarity 21.1%; Pred. No. 31;
Matches 30; Conservative 25; Mismatches 49; Indels 38; Gaps 5;

QY 14 LALLAAPAFVPTVDLLANPAGWISYGNQENYRHSPLTQIT 60
DB 579 LCVVSGHVNAGVAIEHSEIVDEVFYKILMPKFNKNGVPRRMWLSGNCNLESEII 638

QY 61 TENV-----QLOLVARGMQPGKVQVTPPLIHGVMYLANPGDV 99

DB 639 TKWTGSDDLVNTKELAELEKFAEDNELOSEWRKAGNNKMKIVSLIREKGYVSP-DA 697

QY 100 IQAIDAKTGDLIWEHRRPLN 121
DB 698 MEDVQIKR--IHEKROLLNI 716

RESULT 11
US-08-510-215A-2
; Sequence 2, Application US/08510215A
; Patent No. 5814506

GENERAL INFORMATION:
APPLICANT: KONG, Huimin
APPLICANT: PELLETIER, John J.
APPLICANT: ALIOTTA, Jason M.
TITLE OF INVENTION: OVER-EXPRESSION AND PURIFICATION OF
TITLE OF INVENTION: A TRUNCATED THERMOSTABLE DNA POLYMERASE BY PROTEIN
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: GREGORY D. WILLIAMS; NEW ENGLAND
ADDRESSEE: BIOLABS, INC.
STREET: 32 TOZER ROAD
CITY: BEVERLY
STATE: MASSACHUSETTS
COUNTRY: US

ZIP: 01915
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/510,215A
FILING DATE: 02-AUG-1995
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
NAME: WILLIAMS, GREGORY D.
REGISTRATION NUMBER: 30901
REFERENCE/DOCKET NUMBER: NEB-113
TELECOMMUNICATION INFORMATION:
TELEPHONE: (508) 927-5054
TELEFAX: (508) 927-1705

INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 876 amino acids
TYPE: amino acid
TOPOLOGY: linear

MOLECULE TYPE: protein
US-08-510-215A-2

Query Match 10.0%; Score 65.5; DB 2; Length 876;
Best Local Similarity 23.0%; Pred. No. 40;
Matches 38; Conservative 16; Mismatches 54; Indels 57; Gaps 7;

QY 14 LALLAAPA-----FAOVPTVDLLANPAGWISYGNQENYRHSPLTQITTE 62
DB 282 LKMAAPFAEGEKPLEMEFAIVDYITTEMLADKXAA--LVEVMEVNYHDAPIVGIALV 338

QY 63 N-----VGLOLVW-----ARGMQPGKVQVTPLIHGV-----M 91
DB 339 NEHGFEMRPETALADSCFLAMLADETKKSMEDAKRAVVALMKRGIELRGVAFDILLAA 398

QY 92 YLANP-----GDVQAIIDAKTGDLIWEH-----RRPLNATL 124
DB 399 YLLNPAQADADIAVAKMKOYEAVRSDEAVYKGYKRSJPDQTL 443

RESULT 12
US-09-310-293-2
; Sequence 2, Application US/09310293

PAATWYEKHDLSSTDMHPYVHP 796

b 568 LKDLWNDGRIM 578

RESULT 15

```

US-08-852-091-2
; Sequence 2, Application US/08852091
; Patent No. 5883228
; GENERAL INFORMATION:
; APPLICANT: James E. Darnell, Jr.
; APPLICANT: Zilong Wen
; APPLICANT: Curt M. Horvath
; APPLICANT: Zhong Zhong
; TITLE OF INVENTION: FUNCTIONALLY ACTIVE REGIONS OF SIGNAL
; TITLE OF INVENTION: TRANSDUCER AND ACTIVATOR OF TRANSCRIPTION (STAT) PROTEINS
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/852,091
; FILING DATE: 06-MAY-1997
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/369,796
; FILING DATE: 06-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-116
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201 487-5800
; TELEFAX: 201 343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 851 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-852-091-2

```

```

Query Match          9.9%; Score 65; DB 2; Length 851;
Best Local Similarity 22.9%; Pred. No. 44;
Matches 30; Conservative 16; Mismatches 43; Indels 42; Gaps 6;

```

```

QY 1 MKPTSLMASGALALAPAFQVTVUDELANPPAGW-----ISYGO 46
DB 459 MNOISIMASVLAENLSPNLQND-----QFSNPKAPWLLGPAISMOSFSSYVGRGL 512
QY 47 NOENYRHSPLTOITTEVGO-----LQLYVA-----RGMOPKRV-----QVTPLIHDS 89
DB 513 NSDQ-----LSMLRNKLFQGNCRTEPDLISWADFTKRESPPKRLPFTWLDKLTLELVHDH 567
QY 90 VMYTLANPGDVI 100
DB 568 LKDIMNDGRIM 578

```

```

Search completed: May 24, 2002, 10:02:56
Job time: 230 sec

```

10-12

10-13

10-14

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2002, 10:06:02 ; Search time 56.04 Seconds
(without alignments)
214.332 Million cell updates/sec

Title: US-08-934-506a-5_COPY_1_125

Perfect score: 657
Sequence: 1 MKPTSLIMASAGALALIAAP.....KTGDLIWEHRQLPNTATLN 125

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	190	28.9	742	2 A49340	alcohol dehydrogen
2	183	27.9	738	2 S14270	alcohol dehydrogen
3	178	27.1	742	2 JS0326	alcohol dehydrogen
4	165.5	25.2	708	2 AG5217	glucose dehydrogen
5	152.5	23.2	796	2 AG0523	glucose dehydrogen
6	150.5	22.9	796	1 JV0107	glucose dehydrogen
7	150.5	22.9	796	2 H90644	glucose dehydrogen
8	150.5	22.9	796	2 H85495	glucose dehydrogen
9	148	22.5	778	2 G98221	glucose dehydrogen
10	148	22.5	778	2 A13064	glucose dehydrogen
11	137	20.9	623	2 B83399	glucose dehydrogen
12	136.5	20.8	603	2 F83360	alcohol dehydrogen
13	134.5	20.5	601	2 E95863	glucose dehydrogen
14	134	20.4	808	1 QPKEX	glucose dehydrogen
15	116	17.7	801	1 S00943	glucose dehydrogen
16	105.5	16.1	809	2 B98314	probable quinate d
17	105.5	16.1	809	2 JQ0706	hypothetical prote
18	101.5	15.4	626	2 A55547	alcohol dehydrogen
19	94.5	14.4	809	2 JQ4881	quinate-shikimate
20	93.5	14.2	639	2 JC4881	polysialyl-alcohol
21	83.5	12.7	2802	2 F97686	cyclic beta-(1-2)
22	83.5	12.7	2831	2 A12911	beta (1-->2) gluc
23	82	12.5	668	2 C75264	probable serine/th
24	81	12.3	613	2 F69424	conserved hypotet
25	80.5	12.3	647	2 C83012	probable chemotaxi
26	78.5	11.9	148	1 S05638	cuticle protein 8
27	77	11.7	393	2 AD0350	probable 11prote
28	77	11.7	733	2 S78376	photosystem I P700
29	76.5	11.6	400	2 AD0604	D-alanyl-D-alanine

30	76	11.6	306	2 D75562	N-acetyl-gamma-glu
31	76	11.6	386	2 A82284	conserved hypotet
32	76	11.6	573	2 S68591	methanol dehydrog
33	76	11.6	580	2 AB0994	gamma-glutamyltran
34	76	11.6	772	2 H86016	hypothetical prote
35	75.5	11.5	224	2 AD2115	phosphoribosylform
36	75.5	11.5	407	2 C82804	conserved hypotet
37	75.5	11.5	684	2 T36771	probable integral
38	75.5	11.5	3972	2 S75251	hypothetical prote
39	74	11.3	380	2 C83171	conserved hypotet
40	74	11.3	969	2 A75634	McRb-related prote
41	73.5	11.2	584	2 AH2506	ser/thr protein ki
42	73	11.1	355	2 E95890	probable sugar upt
43	73	11.1	392	2 AG0821	probable lipoprote
44	73	11.1	407	2 G87515	secreted protein,
45	72	11.0	244	2 D75505	hypothetical prote

ALIGNMENTS

RESULT 1
A49340
alcohol dehydrogenase (EC 1.1.-.-) precursor - Acetobacter pasteurianus (strain NC113
C:Species: Acetobacter pasteurianus
C>Date: 07-Apr-1994 #sequence-revision 07-Apr-1994 #text-change 20-Jun-2000
C:Accession: A49340
J: Takemura, H.; Kondo, K.; Horiuchi, S.; Beppu, T.
R: Bacteriol. 175, 6857-6866, 1993
A:Title: Induction by ethanol of alcohol dehydrogenase activity in Acetobacter pasteu
A:Reference number: A49340; MUID:94042848
A:Accession: A49340
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-742 <TAK>
A:Cross-references: GB:D13893; NID:g517067; PIDN:BA40252.1; PID:g452586
C:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)
C:Keywords: alcohol metabolism; NAD; oxidoreductase

Query Match 28.9%; Score 190; DB 2; Length 742;
Best Local Similarity 38.8%; Pred. No. 3.2e-11;
Matches 45; Conservative 18; Mismatches 49; Indels 4; Gaps 3;

QY 11 AGALALIAAPNAQVTPYTDLL--ANPPAGEMISYGONOEYRHSPLTQTITENVGOLQ 68
DB 23 AALPYAAYPARADGCGNTGEAIIHADDPHEWLSYGTYSQGRSPIDQINRSWVGLK 82
QY 69 LVMARGMOPGKYQ-VTPLIHDCVMYLANPGDYIAIDAKTGDLIWEHRQLP-NIA 122
DB 83 LAMVYTLDTNRGCEATPLVVDGIMVATTNWMSKEMALDAATGKLWQYDPKVGNTA 138

RESULT 2
S14270
alcohol dehydrogenase, membrane-bound (EC 1.1.-.-) 72K chain precursor - Acetobacter
C:Species: Acetobacter polyoxogenes
A:Variety: strain NB11028
C>Date: 21-Nov-1993 #sequence-revision 26-Jul-1996 #text-change 20-Jun-2000
C:Accession: S14270
R: Tamaki, T.; Fukaya, M.; Takemura, H.; Tayama, K.; Okumura, H.; Kawamura, Y.; Nishiy
Biochim. Biophys. Acta 1088, 292-300, 1991
A:Title: Cloning and sequencing of the gene cluster encoding two subunits of membrane
A:Reference number: S14270; MUID:91159482
A:Accession: S14270
A:Molecule type: DNA
A:Residues: 1-738 <TAM>
A:Cross-references: GB:D00635; NID:g216185; PIDN:BA00528.1; PID:g216186
A:Experimental source: strain NB11028
C:Complex: heterodimer of 72K and 44K chains
C:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)
C:Keywords: alcohol metabolism; membrane protein; oxidoreductase
F:1-35/Domain: signal sequence #status predicted <SIG>

F:36-738/Product: alcohol dehydrogenase 72k chain #status predicted <MAY>

Query Match 27.9%; Score 193; DB 2; Length 738;

Best Local Similarity 33.3%; Pred. No. 1,6e-10;

Matches 41; Conservative 23; Mismatches 49; Indels 10; Gaps 3;

QY 10 SAGALALAAPAFQVTPVDELT-----LLANPPAGEWISYSGONGENRHSPLQTITTT 61
 DB 17 TAGTICALISGATMASADGGATGEGALIHADHPGNMWTGRTYSDDRTSPLDQINR 76
 QY 62 ENVGQQLVWARGMQPGKVQV-TPLIHGVMTLANPGDVIOAIDAKTGLIWEHRQLP- 119
 DB 77 SNVGMKLAWYIDIDFTNGQEGTPLYDGVMTATNWSMKRAVDAATGKLMSYDPRVPG 136
 QY 120 NTA 122
 DB 137 NTA 139

RESULT 3

JS0326

alcohol dehydrogenase (EC 1.1.-.-) 72k chain precursor - Acetobacter aceti

C:Species: Acetobacter aceti

C:Date: 31-Mar-1990 #sequence, revision 31-Mar-1990 #text, change 20-Jun-2000

R:Inoue, T.; Sunagawa, M.; Mori, A.; Imai, C.; Fukuda, M.; Takagi, M.; Yano, K.

J. Bacteriol. 171, 3115-3122, 1989

A:Title: Cloning and sequencing of the gene encoding the 72-kilodalton dehydrogenase sub

A:Reference number: JS0326; MUID:89255070

A:Accession: JS0326

A:Molecule type: DNA

A:Residues: 1-742 <TNC>

A:Cross-references: GB:D90004; GB:M26951; NID:g216193; PIDN:BAA14058.1; PID:g216194

A:Experimental source: strain K6033

A:Note: amino terminal of mature protein is confirmed

C:Genetics:

A:Gene: adh1

C:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)

C:Keywords: alcohol metabolism; NAD; oxidoreductase

F:1.35/Domain: signal sequence #status predicted <SIG>

F:36-742/Product: alcohol dehydrogenase 72k dehydrogenase chain #status predicted <ADC>

Query Match

Best Local Similarity 37.9%; Score 178; DB 2; Length 742;

Matches 44; Conservative 19; Mismatches 49; Indels 4; Gaps 3;

QY 11 AGALALAAPAFQVTPVDELT--ANPPAGEWISYSGONGENRHSPLQTITTEVGOIQ 68
 DB 23 AALPFAAVPARADGGGNGTGEALIHADHPENWLSGRTYSDDRTSPLDQINRNVGDLK 82
 QY 69 LVMARGMQPGKVQ-VTPLIHGVMTLANPGDVIOAIDAKTGLIWEHRQLP-NIA 122
 DB 83 LIGVYTLIDNRGQEGATPLYVDGIMYATNWSMKWALDAATGKLMLQYDKRVGNIA 138

RESULT 4

S52317

quinohemoprotein ethanol dehydrogenase (EC 1.1.99.-) type 1 precursor - Comamonas testos

C:Species: Comamonas testos

C:Date: 08-May-1995 #sequence, revision 21-Jul-1995 #text, change 02-Jun-2000

C:Accession: S62366; S62373; S65908; S52317

R:Stoerovogel, J.; Kraayveld, D.E.; van Sluis, C.A.; Jongejan, J.A.; de Vries, S.; Duine,

Eur. J. Biochem. 235, 690-696, 1996

A:Title: Characterization of the gene encoding quinohemoprotein ethanol dehydrogenase c

A:Reference number: S62366; MUID:96184549

A:Accession: S62366

A:Status: preliminary

A:Molecule type: DNA

A:Cross-references: EMBL:X81880; NID:g663195; PIDN:CAA57464.1; PID:g663196

A:Status: preliminary

A:Molecule type: protein

A:Residues: 32-54 <STO2>

R:de Jong, G.A.H.; Geerlof, A.; Stoerovogel, J.; Jongejan, J.A.; de Vries, S.; Duine,

Eur. J. Biochem. 230, 899-905, 1995

A:Title: Quinohemoprotein ethanol dehydrogenase from Comamonas testosteroni. Purific

A:Reference number: S65908; MUID:95324580

A:Accession: S65908

A:Molecule type: protein

A:Residues: 32-50, 'X', 52-54, 477-483, 'X', 485-490 <DEJ>

A:Experimental source: ATCC 15667

C:Genetics:

A:Gene: qheh

C:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)

C:Keywords: oxidoreductase; quinoprotein

F:1.31/Domain: signal sequence #status predicted <SIG>

F:32-708/Product: quinohemoprotein ethanol dehydrogenase type 1 #status experimental

Query Match

Best Local Similarity 25.2%; Score 165.5; DB 2; Length 708;

Matches 41; Conservative 23; Mismatches 50; Indels 15; Gaps 4;

QY 3 PSLWASAGALALAAPAFQVTPV-----DELANPP-AGEWISYSGONGEN 50
 DB 12 PGRMVLILACLG--SAAAFQGTGPAQAAAVQRYDGFIRANARPPDWPTIGVYAE 69
 QY 51 YRHSPLQTITTEVGOQLVWARGMQPGK-VQVTPLIHGVMTLANPGDVIOAIDAKTGD 109
 DB 70 TRYSRLDQILANAVKDLGLAMSYNLESTRGVATEPVVDGIMYVSAHSVVAIDITRTGN 129
 QY 110 LIWEHRQL 118
 DB 130 RLTWTDPPQI 138

RESULT 5

AG0523

glucose dehydrogenase (imported) - Salmonella enterica subsp. enterica serovar Typh

C:Species: Salmonella enterica subsp. enterica serovar Typh

A:Note: this species has also been called Salmonella typhi

C:Date: 09-Nov-2001 #sequence, revision 09-Nov-2001 #text, change 27-Nov-2001

C:Accession: AG0523

R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Church

th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farr

S.; Moule, S.; O'Gaora, P.

Mature 413, 848-852, 2001

A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens,

A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica se

A:Reference number: AB0502; PMID:11677608

A:Accession: AG0523

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-796 <PAR>

A:Cross-references: GB:AL513382; PIDN:CAD01327.1; PID:g16501455; GSPDB:GNO0176

C:Genetics:

A:Gene: STY0191

C:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)

Query Match 23.2%; Score 152.5; DB 2; Length 796;
 Best Local Similarity 33.3%; Pred. No. 2.3e-07;
 Matches 43; Conservative 17; Mismatches 44; Indels 25; Gaps 5;
 QY 6 LTVMS-----AGALALAAPAFQVTPVDELTANPPAGEWISYSGONGENRHSPLQT 58
 DB 135 LTVMGFNDPQFINTGLTADATPA-ATSSSTADE-----DMPAYGRNDEGRVSPLEK 185
 QY 59 ITTENVGQQLVWA-----RGMQPGKV--QVTPLIHGVMTLANPGDVIOAIDAKTGD 109
 DB 186 ITADNVHQLKAWYFRGDLKQPRNDPEITNEVTPIKVGTGLTCTARHORFALDAASGR 245
 QY 110 LIWEHRQL 118

Db 246 EKMHPDEL 254

RESULT 6

glucose dehydrogenase (pyrroloquinoline-quinone) (EC 1.1.99.17) - Escherichia coli

C:Species: Escherichia coli

C:Date: 30-Jun-1992 #sequence_revision 12-Dec-1997 #text_change 16-Jun-2000

C:Accession: D64735; J0107; A45997; S45201; I41228

R:Blatter, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co

.A.; Rose, D.J.; Mau, B.; Shao, Y.

Science 277, 1453-1462, 1997

A:Title: The complete genome sequence of Escherichia coli K-12.

A:Reference number: A64720; MUID:97426617

A:Accession: D64735

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-796 <BLAT>

A:Cross-references: GB:AE000122; GB:U00096; NID:g1786315; PIDN:AA073235.1; PID:g1786316;

A:Experimental source: strain K-12, substrain M61655

R:Clifton-Jansen, A.M.; Goosen, N.; Fayet, O.; van de Putte, P.

J. Bacteriol. 172, 6308-6315, 1990

A:Title: Cloning, mapping, and sequencing of the gene encoding Escherichia coli quinopri

A:Reference number: J0107; MUID:91035240

A:Accession: J0107

A:Molecule type: DNA

A:Residues: 1-36, 'L', 60-148, 'H', 150, 'KRRCHT', 157-192, 'K', 194-796 <CLE>

A:Cross-references: GB:X51323; NID:941553; PIDN:CA35706.1; PID:941554

A:Experimental source: strain K12

R:Yamada, M.; Sumi, K.; Matsushita, K.; Adachi, O.; Yamada, Y.

J. Biol. Chem. 268, 12812-12817, 1993

A:Title: Topological analysis of quinoprotein glucose dehydrogenase in Escherichia coli

A:Reference number: A45997; MUID:93286127

A:Accession: A45997

A:Status: preliminary

A:Molecule type: protein

A:Residues: 1-20 <YAM>

R:Fujita, N.

submitted to the EMBL Data Library, January 1994

A:Reference number: S45181

A:Accession: S45201

A:Molecule type: DNA

A:Residues: 1-796 <FUJ>

A:Cross-references: EMBL:D26562; NID:g473770; PIDN:BA005580.1; PID:g473791

A:Experimental source: strain K-12 substrain W3110

R:Yamada, M.; Asaka, S.; Saier, M.H.; Yamada, Y.

J. Bacteriol. 175, 568-571, 1993

A:Title: Characterization of the gcd gene from Escherichia coli K-12 W3110 and regulatio

A:Reference number: I41228; MUID:93123180

A:Accession: I41228

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-148, 'H', 150, 'KRRCHT', 157-192, 'K', 194-665, 'H', 667-796 <RES>

A:Cross-references: GB:D12651; NID:g216555; PIDN:BA002174.1; PID:g216556

C:Genetics:

A:Gene: gcd

A:Map position: 3 min

C:Function:

A:Description: catalyzes the oxidation of D-glucose to D-gluconic acid by ubiquinone

A:Pathway: respiratory chain

C:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)

C:Keywords: oxidoreductase; pyrroloquinoline quinone; respiratory chain; transmembrane

F:11-37/Domain: transmembrane #status predicted <TM1>

F:41-59/Domain: transmembrane #status predicted <TM2>

F:63-81/Domain: transmembrane #status predicted <TM3>

F:96-110/Domain: transmembrane #status predicted <TM4>

F:120-140/Domain: transmembrane #status predicted <TM5>

F:93-95/Binding site: ubiquinone (Arg, Asp) #status predicted

F:466/Active site: Asp #status predicted

Query Match 22.9%; Score 150.5; DB 1; Length 796;

Best Local Similarity 31.8%; Pred. No. 3,7e-07; Matches 41; Conservative 20; Mismatches 43; Indels 25; Gaps 5;

Db 135 LTMAS-----AGATALLAFAFAVPTVDLILANPPAGEWISYGOENYRHSPLTQ 58

6 LTMAS-----AGATALLAFAFAVPTVDLILANPPAGEWISYGOENYRHSPLTQ 58

C:Species: Escherichia coli

C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001

C:Accession: H85495

R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; May

iller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apoda

Nature 409, 529-533, 2001

A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.

A:Reference number: A85480; MUID:2107435; PMID:11206551

A:Accession: H85495

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-796 <STO>

A:Cross-references: GB:AE005174; NID:g12512839; PIDN:AA054428.1; GSPDB:GN00145; UMGCP:

A:Experimental source: strain O157:H7, substrain EDL933

C:Genetics:

A:Gene: gdh

A:Map position: 3 min

C:Function:

A:Description: catalyzes the oxidation of D-glucose to D-gluconic acid by ubiquinone

A:Pathway: respiratory chain

C:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)

C:Keywords: oxidoreductase; pyrroloquinoline quinone; respiratory chain; transmembrane

F:11-37/Domain: transmembrane #status predicted <TM1>

F:41-59/Domain: transmembrane #status predicted <TM2>

F:63-81/Domain: transmembrane #status predicted <TM3>

F:96-110/Domain: transmembrane #status predicted <TM4>

F:120-140/Domain: transmembrane #status predicted <TM5>

F:93-95/Binding site: ubiquinone (Arg, Asp) #status predicted

F:466/Active site: Asp #status predicted

Query Match 22.9%; Score 150.5; DB 2; Length 796;

Best Local Similarity 31.8%; Pred. No. 3,7e-07; Matches 41; Conservative 20; Mismatches 43; Indels 25; Gaps 5;

Db 135 LTMAS-----AGATALLAFAFAVPTVDLILANPPAGEWISYGOENYRHSPLTQ 58

6 LTMAS-----AGATALLAFAFAVPTVDLILANPPAGEWISYGOENYRHSPLTQ 58

C:Species: Escherichia coli

C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001

C:Accession: H85495

R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; May

iller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apoda

Nature 409, 529-533, 2001

A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.

A:Reference number: A85480; MUID:2107435; PMID:11206551

A:Accession: H85495

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-796 <STO>

A:Cross-references: GB:AE005174; NID:g12512839; PIDN:AA054428.1; GSPDB:GN00145; UMGCP:

A:Experimental source: strain O157:H7, substrain EDL933

C:Genetics:

A:Gene: gdh

A:Map position: 3 min

C:Function:

A:Description: catalyzes the oxidation of D-glucose to D-gluconic acid by ubiquinone

A:Pathway: respiratory chain

C:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)

C:Keywords: oxidoreductase; pyrroloquinoline quinone; respiratory chain; transmembrane

F:11-37/Domain: transmembrane #status predicted <TM1>

F:41-59/Domain: transmembrane #status predicted <TM2>

F:63-81/Domain: transmembrane #status predicted <TM3>

F:96-110/Domain: transmembrane #status predicted <TM4>

F:120-140/Domain: transmembrane #status predicted <TM5>

F:93-95/Binding site: ubiquinone (Arg, Asp) #status predicted

F:466/Active site: Asp #status predicted

GenCore version 4.5
Copyright (c) 1993 - 2000 Comugen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2002, 10:15:47 ; Search time 27.86 Seconds

(Without alignments)
173.724 Million cell updates/sec

Title: US-08-934-506a-5_COPY_1_125

Perfect score: 657
Sequence: 1 MKPTSLWMSAGALALAAP.....KTGDLMEHRRQLPNIATLN 125

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	ID	Description
1	206	31.4	DHET_GLUOX	O05542 gluconobact
2	183	27.9	DHET_ACEPO	P28036 acetobacter
3	182	27.7	DHET_ACEBU	Q44002 acetobacter
4	178	27.1	DHET_ACEBU	P18278 acetobacter
5	150.5	22.9	DHET_ECOLI	P15877 escherichia
6	137	20.9	EXXA_PSEAE	Q92472 pseudomonas
7	134	20.4	DHG_GLUOX	P27175 gluconobact
8	116	17.7	DHGA_ACTCA	P05465 acetobact
9	112.5	17.1	DHMI_PARDE	P12293 paracoccus
10	111.5	17.0	DHMI_METOR	P15279 methylolact
11	101.5	15.4	DHMI_METEX	P16027 methylolact
12	94.5	14.4	OUIA_ACTCA	O59086 acetobact
13	91	13.9	XOXF_PARDE	P29668 paracoccus
14	90	13.7	OUIA_XANCU	Q5xd78 xanthomonas
15	87.5	11.9	CU08_LOCSI	P11734 locusta mig
16	77	11.7	PSAB_ODOSI	P49480 odontella s
17	76	11.6	TNFA_MARMO	O35734 marmota mon
18	76	11.6	DHMI_METME	P38539 methylolphi
19	73	11.1	PSAB_HETTR	Q9xv27 heterocapsa
20	70.5	10.7	PPBI_RAT	P15693 rattus norv
21	70	10.7	GUN2_THERU	P26222 thermomonos
22	70	10.7	YBHU_ECOLI	P75764 escherichia
23	69.5	10.6	DACC_ECOLI	P08506 escherichia
24	69	10.5	INVI_CAPAN	P33761 capsicum an
25	68.5	10.4	HM2_SELMA	P45623 selaginella
26	68	10.4	YV99_MYCN	P75593 mycoplasma
27	68	10.4	PSAB_PROHO	P58387 prochloroth
28	68	10.4	GLTB_AZOB	Q05735 azospirillum
29	67.5	10.3	AZUP_ALCPA	P04377 alcaligenes
30	67.5	10.3	TPIS_MYCLE	P46711 mycobacteri
31	67.5	10.3	P60_LISIV	O01837 listeria iv
32	67.5	10.3	PSAB_CYACA	O91166 cyanidium c
33	67	10.2	MP63_MYCTU	P91175 mycobacteri

ALIGNMENTS

RESULT	ID	Query Length	STANDARD	PRT	757 AA
1	DHET_GLUOX	31.4	331	1	DIV_ECOLI
AC	O05542	331	1	1	BIND_STERN
DT	01-NOV-1997 (rel. 35, Created)	485	1	1	YEJ1_YEAST
DT	01-NOV-1997 (rel. 35, Last sequence update)	500	1	1	STRA2_PIG
DT	16-OCT-2001 (rel. 40, Last annotation update)	864	1	1	PHS2_SOLIU
DE	Alcohol dehydrogenase [acceptor] precursor (EC 1.1.99.8) (G3-ADH subunit I).	974	1	1	GLTB_SYNY3
DE	ADHA.	66.5	10.1	1	GPDA_XYTRA
OS	Gluconobacter oxydans (Gluconobacter suboxydans).	66	10.0	1	GDPA_XYTRA
OC	Bacteria; Proteobacteria; alpha subdivision; Acetobacteraceae;	66	10.0	1	YDH1_ECOLI
OC	Gluconobacter.	66	10.0	1	YF00_MYCPN
OX	NCBI_Taxid=442;	66	10.0	1	PUR0_SYNY3
RN	[1]	66	10.0	1	RLX3_STAVU
RP	SEQUENCE FROM N.A., AND SEQUENCE OF 36-50.	66	10.0	1	OM51_HAEIN
RC	STRAIN-IFO 12528;	66	10.0	1	
RX	MEDLINE-97208225; PubMed-9055427;	66	10.0	1	
RA	Kondo K., Horiuchi S.;	66	10.0	1	
RT	Characterization of the genes encoding the three-component membrane-bound alcohol dehydrogenase from Gluconobacter suboxydans and their expression in Acetobacter pasteurianus.	66	10.0	1	
RT	Appl. Environ. Microbiol. 63:1131-1138(1997).	66	10.0	1	
RL	-1- FUNCTION: DEHYDRATION OF PRIMARY ALCOHOLS (EXCEPTION: METHANOL) (BY SIMILARITY).	66	10.0	1	
CC	-1- CATALYTIC ACTIVITY: A primary alcohol + acceptor = an aldehyde + reduced acceptor.	66	10.0	1	
CC	-1- COFACTOR: POQ AND HEME.	66	10.0	1	
CC	-1- SUBUNIT: HETEROTRIMER (DEHYDROGENASE, CYTOCHROME AND PROTEIN ADHS), THAT FORMS THE ALCOHOL DEHYDROGENASE COMPLEX.	66	10.0	1	
CC	-1- SUBCELLULAR LOCATION: MEMBRANE-BOUND, FACING THE PERIPLASMIC SPACE (POTENTIAL).	66	10.0	1	
CC	-1- SIMILARITY: BELONGS TO THE BACTERIAL POQ DEHYDROGENASE FAMILY.	66	10.0	1	
CC	THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sdb.ch/announce/ or send an email to license@sdb-sdb.ch).	66	10.0	1	
CC	EMBL: D86375; BAA19753.1; -	66	10.0	1	
DR	HSSP: Q924J7; 1F1G	66	10.0	1	
DR	InterPro: IPR001479; Bac_POQ	66	10.0	1	
DR	InterPro: IPR002372; Bac_POQ_repeat	66	10.0	1	
DR	InterPro: IPR000345; Cyt_heme_bind	66	10.0	1	
DR	Pfam: PF01011; Bacterial_POQ_6	66	10.0	1	
DR	PROSITE: PS00363; BACTERIAL_POQ_1; 1	66	10.0	1	
DR	PROSITE: PS00364; BACTERIAL_POQ_2; 1	66	10.0	1	
DR	PROSITE: PS00190; CYTOCHROME_C_1	66	10.0	1	
KW	Signal	66	10.0	1	
FT	CHAIN	35	757	1	ALCOHOL DEHYDROGENASE [ACCEPTOR]
FT	MOD_RES	35	35	1	PYRROLIDONE CARBOXYLIC ACID
FT	ACT_SITE	342	342	1	BASE (POTENTIAL)

FT BINDING 653 653 HEME (COVALENT) (BY SIMILARITY).
 FT BINDING 656 656 HEME (COVALENT) (BY SIMILARITY).
 FT METAL 657 657 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
 SQ SEQUENCE 757 AA; 82968 MW; 39B9F90E3B947581 CRC64;

Query Match 31.4%; Score 206; DB 1; Length 757;
 Best Local Similarity 43.1%; Pred. No. 9, 8e-13;
 Matches 53; Conservative 19; Mismatches 43; Indels 8; Gaps 5;

OY 7 LMSAGALLALIA--PAPVQV--TPVTELLANPAGEMISYGCNENYRHSPLQIT 61
 Db 16 LLSCAALALFSAVPAVPAQEDTGTATSSDNGGP-GDWLSYGRYSQRYSPLDQINT 74
 OY 62 ENVGOLQVWARGMOPKQVQV-TPLIHGVMTLANPGDVIAIDAKTGDLIEMHRQLP- 119
 Db 75 ENVGKILKAWHYDLDTNRGEGPPLVNGVMTATNWSMKKLDATGATGLMSYDPKVG 134
 OY 120 NIA 122
 Db 135 NIA 137

RESULT 2

DHET_ACEPO STANDARD; PRT; 738 AA.
 AC P28036;

DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)

DE Alcohol dehydrogenase [acceptor] precursor (EC 1.1.99.8).
 GN ADHA.

OS Acetobacter polyoxogenes.
 OC Bacteria; Proteobacteria; alpha subdivision; Acetobacteraceae;
 OC Acetobacter.

OX NCBI_TaxID=439;
 RN [1]

RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC STRAIN=NB11028;
 RA MEDLINE=91159482; PubMed=2001402;
 RA Tamaki T., Fukaya M., Takemura H., Tayama K., Okumura H.,
 RA Kawamura Y., Nishiyama M., Horinouchi S., Bepko T.;
 RT "Cloning and sequencing of the gene cluster encoding two subunits of
 RT membrane-bound alcohol dehydrogenase from Acetobacter polyoxogenes";
 RL Biochim. Biophys. Acta 1088:292-300(1991).

CC -1- FUNCTION: DEHYDRATION OF PRIMARY ALCOHOLS (EXCEPTION: METHANOL).
 CC -1- CATALYTIC ACTIVITY: A primary alcohol + acceptor - an aldehyde +
 CC reduced acceptor.

CC -1- COFACTOR: POQ AND HEME.
 CC -1- SUBUNIT: HETEROTETRAMER (DEHYDROGENASE, CYTOCHROME, AND TWO
 CC SMALLER UNKNOWN SUBUNITS) THAT FORMS THE ALCOHOL DEHYDROGENASE
 CC COMPLEX.

CC -1- SUBCELLULAR LOCATION: MEMBRANE-BOUND, FACING THE PERIPLASMIC
 CC SPACE (POTENTIAL).
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POQ DEHYDROGENASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; D00635; BAA0528.1; -;
 CC PIR; S14270; S14270.
 CC HSSP; Q924J7; 1FIC.

DR InterPro: IPR001479; Bac_POQ.
 DR InterPro: IPR002372; Bac_POQ_repeat.
 DR InterPro: IPR000345; CytC_heme_bind.
 DR Pfam: PF01011; Bacterial_POQ_6.
 DR PROSITE: PS00363; BACTERIAL_POQ_1; 1.
 DR PROSITE: PS00364; BACTERIAL_POQ_2; 1.
 DR PROSITE: PS00364; BACTERIAL_POQ_2; 1.

DR PROSITE: PS00190; CYTOCHROME_C; 1.
 KW Oxidoreductase; POQ; Heme; Periplasmic; Membrane; Signal.
 FT SIGNAL 1 35
 FT CHAIN 36 738
 FT BINDING 650 650 ALCOHOL DEHYDROGENASE [ACCEPTOR].
 FT BINDING 653 653 HEME (COVALENT) (BY SIMILARITY).
 FT METAL 654 654 HEME (COVALENT) (BY SIMILARITY).
 FT METAL 654 654 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
 SQ SEQUENCE 738 AA; 80840 MW; 1E2B6ED7BCD92AF6 CRC64;

Query Match 27.9%; Score 183; DB 1; Length 738;
 Best Local Similarity 33.3%; Pred. No. 1, 8e-10;
 Matches 41; Conservative 23; Mismatches 49; Indels 10; Gaps 3;

OY 10 SAGLALLIALAFAFQVPTVDE-----LLANPAGEMISYGCNENYRHSPLQIT 61
 Db 17 TAGTICAAALISGVTATMSADGCGATGEALIHADHPGNMTYGRYSDDRYSPLDINR 76
 OY 62 ENVGOLQVWARGMOPKQVQV-TPLIHGVMTLANPGDVIAIDAKTGDLIEMHRQLP- 119
 Db 77 SNGVNLKAWHYDLDTNRGEGPPLVNGVMTATNWSMKKAVDAATGATGLMSYDPKVG 136
 OY 120 NIA 122
 Db 137 NIA 139

RESULT 3

DHET_ACEPU STANDARD; PRT; 739 AA.
 AC Q44002; Q07952;

DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)

DE Alcohol dehydrogenase [acceptor] precursor (EC 1.1.99.8).
 GN ADH.

OS Acetobacter europaeus.
 OC Bacteria; Proteobacteria; alpha subdivision; Acetobacteraceae;
 OC Gluconacetobacter.

OX NCBI_TaxID=33995;
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN=DES11 / DSM 6160;
 RA Turner C.A.K.;

RT Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
 RL

CC -1- CATALYTIC ACTIVITY: A primary alcohol + acceptor - an aldehyde +
 CC reduced acceptor.

CC -1- COFACTOR: POQ AND HEME (BY SIMILARITY).
 CC -1- SUBUNIT: HETEROTETRAMER (DEHYDROGENASE, CYTOCHROME, AND TWO
 CC SMALLER UNKNOWN SUBUNITS) THAT FORMS THE ALCOHOL DEHYDROGENASE
 CC COMPLEX (BY SIMILARITY).

CC -1- SUBCELLULAR LOCATION: MEMBRANE-BOUND, FACING THE PERIPLASMIC
 CC SPACE (POTENTIAL).
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POQ DEHYDROGENASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; X82894; CAA58066.1; -;
 CC EMBL; Y09480; CAA70688.1; -;
 CC HSSP; Q924J7; 1FIC.

DR InterPro: IPR001479; Bac_POQ.
 DR InterPro: IPR002372; Bac_POQ_repeat.
 DR InterPro: IPR000345; CytC_heme_bind.
 DR Pfam: PF01011; Bacterial_POQ_6.
 DR PROSITE: PS00363; BACTERIAL_POQ_1; 1.
 DR PROSITE: PS00364; BACTERIAL_POQ_2; 1.
 DR PROSITE: PS00190; CYTOCHROME_C; 1.

RA Blatner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of *Escherichia coli* K-12.";
 RL Science 277:1453-1474(1997).
 RN [5]
 RP TOPOLOGY.
 RX MEDLINE-93286127; PubMed-8509415;
 RA Yamada M., Sumi K., Matsushita K., Adachi O., Yamada Y.;
 RT "Topological analysis of quinoprotein glucose dehydrogenase in
 RT *Escherichia coli* and its ubiquinone-binding site.";
 RL J. Biol. Chem. 268:12812-12817(1993).
 RN [6]
 RP 3D-STRUCTURE MODELING.
 RX MEDLINE-96128046; PubMed-8554505;
 RA Cozier G.E., Anthony C.;
 RT "Structure of the quinoprotein glucose dehydrogenase of *Escherichia*
 RT *extorquens*.";
 RL Biochem. J. 312:679-685(1995).
 CC -1- FUNCTION: GDH IS PROBABLY INVOLVED IN ENERGY CONSERVATION RATHER
 CC THAN IN SUGAR METABOLISM.
 CC -1- CATALYTIC ACTIVITY: D-glucose + acceptor = D-glucono-1,5-lactone +
 CC reduced acceptor.
 CC -1- COFACTOR: POQ.
 CC -1- SUBUNIT: MONOMER.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane.
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POQ DEHYDROGENASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X51323; CA35706.1; ;
 DR EMBL: D12851; BA02174.1; ;
 DR EMBL: D26562; CAB20298.1; ;
 DR EMBL: AE000122; AAC73235.1; ;
 DR PIR: JVO107; JVO107;
 DR HSSP: P38539; 4AAH;
 DR EcGene: EG10369; gcd.
 DR InterPro: IPR001478; Bac_PQO.
 DR InterPro: IPR002372; Bac_PQO_repeat.
 DR Pfam: PF01011; Bacterial_POQ_7.
 DR PROSITE: PS00363; BACTERIAL_POQ_1; 1.
 DR PROSITE: PS00364; BACTERIAL_POQ_2; 1.
 KW Oxidoreductase; POQ; Transmembrane; Inner membrane; Periplasmic;
 KW Complete proteome.
 FT DOMAIN 1 10
 FT TRANSMEM 11 10
 FT TRANSMEM 37 37
 FT DOMAIN 38 40
 FT TRANSMEM 41 58
 FT DOMAIN 59 62
 FT TRANSMEM 63 81
 FT DOMAIN 82 95
 FT TRANSMEM 96 110
 FT DOMAIN 111 118
 FT TRANSMEM 119 141
 FT DOMAIN 142 142
 FT ACT SITE 466
 FT CONFLICT 59 59
 FT CONFLICT 149 156
 FT CONFLICT 193 193
 FT CONFLICT 666 666
 SQ SEQUENCE 796 AA; 86747 MW; D9EDC705A12894E9 CRC64;
 Query Match 22.9%; Score 150.5; DB 1; Length 796;
 Best Local Similarity 31.8%; Pred. No. 3.1e-07;

Matches 41: Conservative 20: Mismatches 43: Indels 25: Gaps 5:
 QY 6 LIMS-----ACALLAAPAQVPTWDELLANPPAGEWISGQNDNRHSPLTQ 58
 DB 135 LTMAGFNDPDRINLTSLADATPAEA-ISPVDQ-----DMPAYGRNGEGRFSLKQ 185
 QY 59 ITTENVGOLALVMA-----RGMPGKV--QVPELHIDGVMTYANGDVYQAIADAKTD 109
 DB 186 INADNHNKAKAWFRFGDVAKQPNPDEITNEVPIKVGDTLYCTAHRFLDALDASGK 245
 QY 110 LIMEHRQL 118
 DB 246 EKWHYDEPL 254
 RESULT 6
 EXAA_PSEAE STANDARD; PRT; 623 AA.
 ID EXAA_PSEAE
 AC Q924J7;
 DT 30-MAY-2000 (Rel. 39, Created)
 DF 30-MAY-2000 (Rel. 39, Last sequence update)
 DE 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Quinoprotein ethanol dehydrogenase precursor (EC 1.1.99.-) (OEDH).
 GN EXAA OR PA1982.
 OS Pseudomonas aeruginosa.
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 OC Pseudomonas
 OX NCBI_TaxID:287;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ATCC 17933;
 RX MEDLINE-99041560; PubMed-9826187;
 RA Diehl A., Wintzingerode F., Goerisch H.;
 RT "Quinoprotein ethanol dehydrogenase of *Pseudomonas aeruginosa* is a
 RT homodimer: Sequence of the gene and deduced structural properties of
 RT the enzyme.";
 RL Eur. J. Biochem. 257:409-419(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ATCC 15692 / PA01;
 RX MEDLINE-20437337; PubMed-10984043;
 RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warrenner P.,
 RA Hickey M.J., Brinkman F.S.L., Huftagle W.O., Kowalik D.J., Lagrou M.,
 RA Garber R.L., Goltzy L., Tolentino E., Westbrook-Madson S., Yuan Y.,
 RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
 RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
 RA Reiser J., Saler M.H., Hancock R.E.W., Lory S., Olson M.V.;
 RT "Complete genome sequence of *Pseudomonas aeruginosa* PA01, an
 RT opportunistic pathogen.";
 RL Nature 406:959-964(2000).
 RN [3]
 RP SEQUENCE OF 1-52 FROM N.A.
 RC STRAIN-ATCC 17933;
 RX MEDLINE-99173751; PubMed-10075429;
 RA Schober M., Goerisch H.;
 RT "Cytochrome c550 is an essential component of the quinoprotein ethanol
 RT oxidation system in *Pseudomonas aeruginosa*: cloning and sequencing of
 RT the genes encoding cytochrome c550 and an adjacent acetaldehyde
 RT dehydrogenase.";
 RL Microbiology 145:471-481(1999).
 RN [4]
 RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF 35-616.
 RC STRAIN-ATCC 17933;
 RX MEDLINE-20202376; PubMed-10736230;
 RA Keitel T., Diehl A., Knaute T., Stezowski J.J., Hoehne W.,
 RA Goerisch H.;
 RT "X-ray structure of the quinoprotein ethanol dehydrogenase from
 RT *Pseudomonas aeruginosa*: basis of substrate specificity.";
 RL J. Mol. Biol. 297:961-974(2000).
 CC -1- FUNCTION: OXIDIZE PRIMARY ALCOHOLS AND ALSO ACTS ON SECONDARY
 CC ALCOHOL, BUT NOT HIGHLY ACTIVE ON METHANOL.
 CC -1- COFACTOR: POQ AND CALCIUM.
 CC -1- SUBUNIT: HOMODIMER.

CC -1- SUBCELLULAR LOCATION: periplasmic.
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POO DEHYDROGENASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AJ009858; CA00886.1; -
 CC EMBL: AE004624; AAC05370.1; -
 CC EMBL: AF068264; AAC79657.1; -
 CC PDB: 1FG; 30-AUG-00.
 CC InterPro: IPR001479; Bac_POO.
 CC InterPro: IPR002372; Bac_POO_repeat.
 CC Pfam: PF01011; Bacterial_POO_7.
 CC PROSITE: PS00363; BACTERIAL_POO_1; FALSE_NEG.
 CC PROSITE: PS00364; BACTERIAL_POO_2; FALSE_NEG.
 CC Oxidoreductase; POO; Periplasmic; Signal; Calcium; 3D-structure;
 CC Complete proteome.
 CC SIGNAL 1 34
 CC CHAIN 139 623 QUINOPROTEIN ETHANOL DEHYDROGENASE
 CC DISULFID 139 140
 CC SEQUENCE 623 AA; 68123 MW; 32DDEDF20B291D6 CRC64;
 SQ
 Query Match 20.9%; Score 137; DB 1; Length 623;
 Best Local Similarity 31.5%; Pred. No. 5, 1e-06;
 Matches 40; Conservative 20; Mismatches 57; Indels 10; Gaps 4;
 QY 3 PSLIMASAGAL---LAAPAFQVTPYTDLLA--PPAGMISYSGNNGNRYHSPLT 57
 DB 9 PAGILRPSLHCLFAVALGAGALAKDVWEIDANDKTGVDVLYGKTHAKRSPKL 68
 QY 58 QITTEENYGOLVWARGM---OPGKQVYPLIHGVMYLANPGDYIAIDAKTGLDLE 113
 DB 69 QVANADNFKILPMASYSFGDEKORGO-ESQAIYSDGYIYATASYSRLFALDAKTGRRLMT 127
 QY 114 HRQLPN 120
 DB 128 YNHRLPD 134
 RESULT 7
 DHG_GLUOX STANDARD; PRT; 808 AA.
 ID DHG_GLUOX
 AC P27175;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Glucose dehydrogenase [pyrroloquinoline-quinone] precursor
 DE (EC 1.1.99.17).
 GN GDR.
 OS Gluconobacter oxydans (Gluconobacter suboxydans).
 CC Bacteria; Proteobacteria; alpha subdivision; Acetobacteraceae;
 CC Gluconobacter.
 CC NCBI_TaxID=442;
 RN [1]
 RN SEQUENCE FROM N.A.
 RN MEDLINE=92017653; PubMed=1833618;
 RA Cleton-Jansen A.-M., Dekker S., van de Putte P., Goosen N.;
 RT "A single amino acid substitution changes the substrate specificity
 RT of quinoprotein glucose dehydrogenase in Gluconobacter oxydans";
 RL Mol. Gen. Genet. 229:206-212(1991).
 RN [2]
 RN REVISION TO 213.
 RA Goosen N.;
 RA Submitted (FEB-1992) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: D-glucose + acceptor = D-glucono-1,5-lactone +
 CC reduced acceptor.
 CC -1- Cofactor: POO.

CC -1- SUBCELLULAR LOCATION: OUTSIDE OF THE PERIPLASMIC MEMBRANE.
 CC -1- MISCELLANEOUS: THE P1 FORM CAN OXIDIZE ONLY D-GLUCOSE, WHILE THE
 CC P2 FORM CAN ALSO OXIDIZE MALTOSE. THE SEQUENCE OF P1 FORM IS SHOWN
 CC HERE.
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POO DEHYDROGENASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: X62710; CAA44594.1; ALT_SEQ.
 CC PIR: S17716; OPRX.
 CC InterPro: IPR001479; Bac_POO.
 CC InterPro: IPR002372; Bac_POO_repeat.
 CC Pfam: PF01011; Bacterial_POO_7.
 CC PROSITE: PS00363; BACTERIAL_POO_1; 1.
 CC PROSITE: PS00364; BACTERIAL_POO_2; 1.
 CC Oxidoreductase; POO; Transmembrane; Periplasmic; Signal.
 CC SIGNAL 1 33
 CC CHAIN 34 808 GLUCOSE DEHYDROGENASE [PYRROLOQUINOLINE-
 CC TRANSMEM 35 54 POTENTIAL.
 CC TRANSMEM 59 76 POTENTIAL.
 CC TRANSMEM 94 108 POTENTIAL.
 CC TRANSMEM 123 138 POTENTIAL.
 CC ACT SITE 470 470 BASE (POTENTIAL).
 CC VARIANT 788 788 H -> N (IN P2 FORM).
 CC SEQUENCE 808 AA; 87567 MW; 0F4160DA78652445 CRC64;
 SQ
 Query Match 20.4%; Score 134; DB 1; Length 808;
 Best Local Similarity 27.8%; Pred. No. 1, 3e-05;
 Matches 35; Conservative 20; Mismatches 51; Indels 20; Gaps 3;
 QY 13 ALALLAAPFAQVTPYTDLLA-----NPPAGMISYSGNNGNRYHSPLTQITTT 61
 DB 128 AVIALFASLFTDPHDISGLPTQIANASPADPDNYPASMAHAYGTQAGDRSPINQNA 187
 QY 62 ENVGOLVW-----ARGMQPKV--QVTPLIHGVYLANPGDYIAIDAKTGLDLE 112
 DB 188 TIVSNLKVAMHHTKDMNMSNDPGEQTEATPILFEFNTLYMCSLHKLFAYDGAATGNVKK 247
 QY 113 EHRRL 118
 DB 248 VYDPKL 253
 RESULT 8
 DHGA_ACICA STANDARD; PRT; 801 AA.
 ID DHGA_ACICA
 AC P05465;
 DT 01-NOV-1988 (Rel. 09, Created)
 DT 01-NOV-1988 (Rel. 09, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Glucose dehydrogenase-A [pyrroloquinoline-quinone] precursor
 DE (EC 1.1.99.17) (Quinoprotein glucose DH) (GDH-A).
 GN GDHA.
 OS Acinetobacter calcoaceticus.
 CC Bacteria; Proteobacteria; gamma subdivision; Moraxellaceae;
 CC Acinetobacter.
 CC NCBI_TaxID=471;
 RN [1]
 RN SEQUENCE FROM N.A.
 RN STRAIN=LMD 79.41;
 RC MEDLINE=88289368; PubMed=3399393;
 RA Cleton-Jansen A.-M., Goosen N., Odle G., van de Putte P.;
 RT "Nucleotide sequence of the gene coding for quinoprotein glucose
 RT dehydrogenase from Acinetobacter calcoaceticus";
 RL Nucleic Acids Res. 16:6228-6228(1988).

```

CC -1- CATALYTIC ACTIVITY: D-glucose + acceptor = D-glucono-1,5-lactone +
CC reduced acceptor.
CC -1- COFACTOR: POQ.
CC -1- SUBUNIT: MONOMER.
CC -1- SUBCELLULAR LOCATION: Periplasmic.
CC -1- MISCELLANEOUS: ACINETOBACTER CALCOACETICUS CONTAINS TWO DIFFERENT
CC POQ DEPENDENT ENZYMES WITH GDH ACTIVITY. GDH-A PREFERS
CC 2-DEOXYGLUCOSE AS SUBSTRATE, THE SPECIFIC SUBSTRATES FOR GDH-B ARE
CC DISACCHARIDES.
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POQ DEHYDROGENASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X07235; CAA30222.1; -.
DR PIR: S00943; S00943.
DR InterPro: IPR001479; Bac_POQ.
DR InterPro: IPR002372; Bac_POQ_repeat.
DR Pfam: PF01011; Bacterial_POQ_7.
DR PROSITE: PS00363; BACTERIAL_POQ_1; 1.
DR PROSITE: PS00364; BACTERIAL_POQ_2; 1.
KW Oxidoreductase; POQ; Transmembrane; Periplasmic; Signal.
FT SIGNAL 1 33
FT CHAIN 34 801
FT TRANSMEM 39 55 [PYROLOQUINOLINE-QUINONE].
FT TRANSMEM 59 79 POTENTIAL.
FT TRANSMEM 94 108 POTENTIAL.
FT TRANSMEM 119 138 POTENTIAL.
FT ACT_SITE 471 471 BASE (POTENTIAL).
SQ SEQUENCE 801 AA; 86956 MW; 2F4EA222EB5450D8 CRC64;

Query Match 17.7%; Score 116; DB 1; Length 801;
Best Local Similarity 29.8%; Pred. No. 0.0008;
Matches 31; Conservative 14; Mismatches 39; Indels 20; Gaps 3;

QY 20 PAFQAQVPTVDLLANPPAGEWISYSGONENYRHSPLTQTTEYNQOLOLWARGKOPK 79
DB 152 PETAQAAPGVAAE-----SDMPAYGRITQGVYRSPKLOINDQNVADLKAAMT--LRTGD 202
QY 80 V-----QVTPLIHGVVYLANPGVYQAIADKAGDLIW 112
DB 203 LKTDNDSGETTNOVPTIKIGNNMFICTAHQOLIAIDPATGKEKW 246

RESULT 9
DHML_METOR STANDARD; PRT; 631 AA.
ID DHML_METOR
AC P12293;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Methanol dehydrogenase subunit 1 precursor (EC 1.1.99.8) (MDH large
DE alpha subunit) (MEDH).
GN MOXF.
OS Paracoccus denitrificans.
OC Bacteria; Proteobacteria; alpha subdivision; Rhodobacter group;
OC Paracoccus.
OX NCBI_TaxID=266;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 33-49.
RX MEDLINE=87307969; PubMed=3114231;
RA Harms N., de Vries G.E., Maurer K., Hoogendijk J., Stoutamer A.H.;
RT "Isolation and nucleotide sequence of the methanol dehydrogenase
RT structural gene from Paracoccus denitrificans.";
RL J. Bacteriol. 169:3969-3975 (1987).
-1- CATALYTIC ACTIVITY: A primary alcohol + acceptor = an aldehyde +

```

```

CC -1- reduced acceptor.
CC -1- COFACTOR: POQ.
CC -1- SUBCELLULAR LOCATION: Periplasmic.
CC -1- MISCELLANEOUS: MDH IS THE MAJOR PROTEIN IN THE CELL DURING GROWTH
CC ON METHANOL (IN P. DENITRIFICANS MDH CONSTITUTES UP TO 15% OF THE
CC TOTAL CELL PROTEIN).
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POQ DEHYDROGENASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M17339; AAA88366.1; -.
DR HSSP: P38539; 4AAH.
DR InterPro: IPR001479; Bac_POQ.
DR InterPro: IPR002372; Bac_POQ_repeat.
DR Pfam: PF01011; Bacterial_POQ_7.
DR PROSITE: PS00363; BACTERIAL_POQ_1; 1.
DR PROSITE: PS00364; BACTERIAL_POQ_2; 1.
KW Oxidoreductase; POQ; Signal; Methanol utilization; Periplasmic.
FT SIGNAL 1 32
FT CHAIN 33 631
FT DISULFID 135 136 METHANOL DEHYDROGENASE SUBUNIT 1.
FT DISULFID 418 447 BY SIMILARITY.
FT ACT_SITE 335 335 BASE (POTENTIAL).
SQ SEQUENCE 631 AA; 69799 MW; 0934DC93FEC57308 CRC64;

Query Match 17.1%; Score 112.5; DB 1; Length 631;
Best Local Similarity 32.6%; Pred. No. 0.0014;
Matches 42; Conservative 25; Mismatches 47; Indels 15; Gaps 9;

QY 4 TSLMASAGALLT-AAFAQVPTVDL--LANPPAGEWISYSGONENYRHSPLTQT 60
DB 12 SLSLMAVAMGLAVLTAPATA-----NDQIVELANDPA-NMYMGRDINAQNSMDIDIN 65
QY 61 TENVOLOLYV--ANGMDGKQVPLPIHGVVYLANP-GVYQAIAD-KTGDLIWEHR- 115
DB 66 KENVVQLRPMSEFSGVVGHG-EGTPLYGVDRMFHTFPPTTALDLNFGKILMOKR 124
QY 116 RQIPNIATL 124
DB 125 KQNPATATV 133

RESULT 10
DHML_METOR STANDARD; PRT; 626 AA.
ID DHML_METOR
AC P15279;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Methanol dehydrogenase subunit 1 precursor (EC 1.1.99.8) (MDH large
DE alpha subunit) (MEDH).
GN MOXF.
OS Methylobacterium organophilum XX.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Methylobacterium group; Methylobacterium.
OX NCBI_TaxID=410;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 29-43.
RX STRAIN=ATCC 27886 / DSM 760 / NCIB 11278;
RX MEDLINE=89008094; PubMed=2459109;
RA Machlin S.M., Hanson R.S.;
RT "Nucleotide sequence and transcriptional start site of the
RT methylobacterium organophilum XX methanol dehydrogenase structural
RT gene.";
RL J. Bacteriol. 170:4739-4747 (1988).
-1- CATALYTIC ACTIVITY: A primary alcohol + acceptor = an aldehyde +

```

```

CC      reduced acceptor.
CC      -1- COFACTOR: POO.
CC      -1- SUBUNIT: THE HOLOENZYME MDH HAS AN ALPHA-2/BETA-2 CONFIGURATION
CC      OF A LARGE ALPHA AND A SMALL BETA SUBUNIT.
CC      -1- SUBCELLULAR LOCATION: PERIPLASMIC, BUT ASSOCIATED WITH INNER
CC      MEMBRANE.
CC      -1- SIMILARITY: BELONGS TO THE BACTERIAL POO DEHYDROGENASE FAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC      or send an email to license@sib-sib.ch).
CC      -----
DR      EMBL; M22629; AAA50289.1; -.
DR      HSSP; P38539; AAAB.
DR      InterPro; IPR001479; Bac_POO.
DR      InterPro; IPR002372; Bac_POO_repeat.
DR      Pfam; PF01011; Bacterial_POO_7.
DR      PROSITE; PS00363; BACTERIAL_POO_1; 1.
DR      PROSITE; PS00364; BACTERIAL_POO_2; 1.
KW      Oxidoreductase; POO; Signal; Methanol utilization; Periplasmic.
FT      SIGNAL 1 28
FT      CHAIN 29 626 METHANOL DEHYDROGENASE SUBUNIT 1.
FT      DISULFID 130 131 BY SIMILARITY.
FT      DISULFID 413 442 BY SIMILARITY.
FT      ACT_SITE 330 330 BASE (POTENTIAL).
SQ      SEQUENCE 626 AA; 68677 MW; 8768FEB8371E5DF CRC64;

Query Match 17.0%; Score 111.5; DB 1; Length 626;
Best Local Similarity 23.2%; Pred. No. 0.0017;
Matches 39; Conservative 26; Mismatches 38; Indels 65; Gaps 7;

OY      10 SAGALALLA-AP-APAQTPVTDELLANPRAE-WISGQNGENRHSPLQITENQO 66
        | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      7 SVSALAMALLAALASSVAYANDKIVELSSDDNNVMPGKNDSNNYSPLKVNKSNVQ 66
OY      67 IQLVW-ARGMQPGVQVYTPLIHDGVMY-----LANGDV----- 99
        | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      67 LRPAMTFSTGLNGH-EGAPLYVDGSKMIVHYSFPNNTRALDDDDGHIIIMQKPNRA 125
OY      100 -----IQAIDAKTGTLDE 113
        : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      126 RAVACCDLVNRLAYWPGDGKTPALILKTYQDRHVVALNAETGETVWK 173

RESULT 11
DHM1_METEX DHM1_METEX STANDARD; PRT; 626 AA.
AC          P16027;
DT      01-APR-1990 (Rel. 14, Created)
DT      01-APR-1990 (Rel. 14, Last sequence update)
DT      15-JUN-1999 (Rel. 38, Last annotation update)
DE      Methanol dehydrogenase subunit 1 precursor (EC 1.1.99.8) (MDH large
EN      alpha subunit) (MDH).
GN      MOXF.
OS      Methylobacterium extorquens.
OC      Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC      Methylobacterium group; Methylobacterium.
OX      NCBI_TaxID=408;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=AM1 / NCIB 9133;
RX      MEDLINE=90337342; PubMed=2116368;
RA      Anderson D.J., Morris C.J., Nunn D.N., Anthony C., Lidstrom M.E.;
RT      "Nucleotide sequence of the methylobacterium extorquens AM1 moxf and
RL      mox genes involved in methanol oxidation."
RN      [2]
RP      SEQUENCE OF 28-53.

```

```

RC STRAIN-AM1 / NCIB 9133;
RX MEDLINE=89350892; PubMed=2504152;
RA Nunn D.N., Day D., Anthony C.;
RT "The second subunit of methanol dehydrogenase of Methylobacterium
RL extorquens AM1."
RL Biochem. J. 260:857-862(1989).
RN [3]
RP DISULFIDE BONDS.
RX MEDLINE=95384759; PubMed=7656012;
RA Blake C.C.F., Ghosh M., Harlos K., Avezoux A., Anthony C.;
RT "The active site of methanol dehydrogenase contains a disulphide
RL bridge between adjacent cysteine residues."
RN Nat. Struct. Biol. 1:102-105(1994).
RN [4]
RP X-RAY CRYSTALLOGRAPHY (1.94 ANGSTROMS).
RX MEDLINE=95253818; PubMed=7735834;
RA Ghosh M., Anthony C., Harlos K., Goodwin M.G., Blake C.;
RT "The refined structure of the quinoprotein methanol dehydrogenase
RL from Methylobacterium extorquens at 1.94 A."
RL Structure 3:177-187(1995).
CC -I- CATALYTIC ACTIVITY: A primary alcohol + acceptor = an aldehyde +
CC reduced acceptor.
CC -I- COFACTOR: PO4.
CC -I- SUBUNIT: THE HOEENZYMME MDH HAS AN ALPHA-2/BETA-2 CONFIGURATION
CC OF A LARGE ALPHA AND A SMALL BETA SUBUNIT.
CC -I- SUBCELLULAR LOCATION: PERIPLASMIC, BUT ASSOCIATED WITH INNER
CC MEMBRANE.
CC -I- SIMILARITY: BELONGS TO THE BACTERIAL POQ DEHYDROGENASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL: M31108; AA25380.1; -.
CC DR PIR: S07908; S07908.
CC DR PIR: JQ0706; JQ0706.
CC DR HSSP: P38539; AAH.
CC DR InterPro: IPR001479; Bac_POQ.
CC DR InterPro: IPR002372; Bac_POQ_repeat.
CC DR Pfam: PF01011; Bacterial_POQ_7.
CC DR PROSITE: PS00363; BACTERIAL_POQ_1; 1.
CC DR PROSITE: PS00364; BACTERIAL_POQ_2; 1.
CC KW Oxidoreductase; POQ; Signal; Methanol utilization; Periplasmic.
CC FT SIGNAL 1 27
CC FT CHAIN 28 626 METHANOL DEHYDROGENASE SUBUNIT 1.
CC FT DISULFID 130 131
CC FT DISULFID 413 442
CC FT ACT_SITE 330 330 BASE (POTENTIAL).
CC SQ SEQUENCE 626 AA; 68434 MW; 64988BD0ARD2AD34C CRC64;

```

ID	QUINA ACICA	STANDARD:	PRT:	809 AA.
AC	Q59086;			
BT	16-OCT-2001 (Rel. 40, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Quinate/shikimate dehydrogenase [Pyruvate:quinone]			
DE	(EC 1.1.99.25) (NAD(P)-independent quinate dehydrogenase).			
GN	QUINA.			
OS	Acinetobacter calcoaceticus.			
OC	Bacteria; Proteobacteria; gamma subdivision; Moraxellaceae;			
OC	Acinetobacter.			
OX	NCBI_TaxID=471;			
RP	[1]			
RP	SEQUENCE FROM N.A., AND CHARACTERIZATION.			
RC	STRAIN-BD413 / ADP1.			
RC	MEDLINE=95095936; PubMed=8002591;			
RA	Elsmore D.A., Ornstom L.N.;			
RA	"The poa-pob supragenomic cluster of Acinetobacter calcoaceticus			
RT	contains quia, the structural gene for quinate-shikimate			
RT	dehydrogenase.";			
RL	J. Bacteriol. 176:7659-7666(1994).			
RN	[2]			
RP	SEQUENCE OF 1-20 FROM N.A.			
RC	STRAIN-BD413 / ADP1.			
RC	MEDLINE=96011389; PubMed=7592351;			
RA	Elsmore D.A., Ornstom L.N.;			
RA	"Unusual ancestry of dehydrogenases associated with quinate catabolism			
RT	in Acinetobacter calcoaceticus.";			
RL	J. Bacteriol. 177:5971-5978(1995).			
CC	-1- FUNCTION: CAN ACT EITHER ON QUINATE OR ON SHIKIMATE.			
CC	-1- CATALYTIC ACTIVITY: quinate + pyruvate:quinone = 5-			
CC	dehydroquininate + reduced pyruvate:quinone.			
CC	-1- CATALYTIC ACTIVITY: SHIKIMATE + PYRUVATE:QUINONE = 3-			
CC	DEHYDROSHIKIMATE + REDUCED PYRUVATE:QUINONE.			
CC	-1- COFACTOR: POO.			
CC	-1- PATHWAY: QUINIC ACID CATABOLIC PATHWAY; FIRST STEP. THIS PATHWAY			
CC	ALLOWS GROWTH OF BACTERIA WITH QUINATE BY ITS CONVERSION TO			
CC	PYRUVATE. QUINATE AND SUBSEQUENT METABOLISM BY THE BETA-KETOGLUTARATE			
CC	PATHWAY.			
CC	-1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).			
CC	-1- INDUCTION: BY PYRUVATE.			
CC	-1- SIMILARITY: BELONGS TO THE BACTERIAL POO DEHYDROGENASE FAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL, L05770; AAC37161.1; -.			
DR	InterPro; IPR001479; Bac_POO.			
DR	InterPro; IPR002372; Bac_POO_repeat.			
DR	Pfam; PF01011; Bacterial_POO_7.			
DR	PROSITE; PS00363; BACTERIAL_POO_2; FALSE_NEG.			
DR	PROSITE; PS00364; BACTERIAL_POO_2; FALSE_NEG.			
KW	Oxidoreductase; POO; Quinate metabolism; Transmembrane.			
FT	TRANSMEM 14 34			
FT	POTENTIAL.			
FT	TRANSMEM 41 61			
FT	POTENTIAL.			
FT	TRANSMEM 68 88			
FT	POTENTIAL.			
FT	TRANSMEM 90 110			
FT	POTENTIAL.			
FT	TRANSMEM 127 147			
FT	POTENTIAL.			
QO	SEQUENCE 809 AA; 88196 MW; 7167CEBBA62BFCB CRC64;			

Query Match	14.4%	Score 94.5;	DB 1;	Length 809;
Best Local Similarity	29.2%;	Pred. No. 0.11;		
Matches 35;	Conservative 13;	Mismatches 49;	Indels 23;	Gaps 4;
OY	3	PTSLIMASGALATLAA-PARAGVPPVIDELIANPAGEWISYGCONQENRHSPLQTITT	61	
I	:	I I :	I I :	I I I

Df 148 PHEIYKASEEELPLVPVDPARKQY-----NMHDYGNDAGSGRFVALDQLNR 193

Oy 62 EVNGOLQLVW-----ARGMOPG-KVOYTPIIHGWATLANPGVIAIDAKTGDLIWE 113
|| : || : | : || : || : || : || : || : || : || : || : || : || : || : || :

Dd 194 NNVSRLKEAWMRPRTDFTTGTGNAGAEDOMTPLOVGNNKYFLCTPHNHNIFAIDADSDGRQLWK 253

RESULT 13

ID	XOXF_PARDE	STANDARD:	PRT.	600 AA.
AC	P29968;			
Dt	01-APR-1993 (Rel. 25, Created)			
Dt	01-OCT-1996 (Rel. 34, Last sequence update)			
Dt	15-DEC-1998 (Rel. 37, Last annotation update)			
De	Putative dehydrogenase XOXF precursor (C1.1.1.99.-).			
Gn	XOXF.			
Os	Paracoccus denitrificans.			
Oc	Bacteria; Proteobacteria; alpha subdivision; Rhodobacter group; Paracoccus.			
Ox	NCBL_TaxID=266;			
Rn	[1]			
Ra	Harms N.; Submitted (AUG-1995) to the EMBL/Genbank/DDBJ databases. [2]			
Rn	SEQUENCE OF 380-600 FROM N.A.			
Rc	STRAIN-PD 1235;			
Rx	MEDLINE-92041583; PubMed-1657873;			
Ra	Ras J., Reijnders W.N.M., van Spanning R.J.M., Harms N., Oltmann L.F., Stouthamer A.H.;			
Rt	"Isolation, sequencing, and mutagenesis of the gene encoding cytochrome c553l of Paracoccus denitrificans and characterization of the mutant strain".			
Rl	J. Bacteriol. 173:6971-6979(1991).			
Cc	- COFACTOR: POO (BY SIMILARITY).			
Cc	- SIMILARITY: BELONGS TO THE BACTERIAL POO DEHYDROGENASE FAMILY.			
Cc	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).			
Cc	EMBL: U34346; AAC4455.1; .			
Dd	EMBL: M75583; AAA25574.1; .			
Dd	PIR: A41378; A41378.			
Dd	HSSP: P38539; AAAH.			
Dd	InterPro: IPR002372; Bac_POO_repeat.			
Dd	Pfam: PF01011; Bacterial_POO; 7.			
Kw	Oxidoreductase; POO; Signal.			
Ft	SIGNAL 1 21 POTENTIAL.			
Ft	CHAIN 22 600 PUTATIVE DEHYDROGENASE XOXF.			
Ft	ACT_SITE 318 318 BASE (POTENTIAL).			
So	SEQUENCE 600 AA; 65159 MW; DCA936FEIBC5A3CE CRC64;			

```

Query Match          13.9%; Score 91; DB 1; Length 600;
Best Local Similarity 28.2%; Pred. No. 0.17;
Matches 35; Conservative 22; Mismatches 45; Indels 22; Gaps 7;

QY 10 SAGATALLAAPFAQVTEVTDELLANP-----PAGEMISYSGQENYRHSPLTOITTE 62
    : |||| : |||| : |||| : |||| : |||| : |||| : |||| : |||| :
Db 6 NGACIALLMSTGTA-----LANEQRAGRDRQAPQAIQMGDYATNRYSTLDQINKD 56

QY 63 NVGQLQLYW--ARKMQPGKVQVTPLIHDGWMYLANP-GDVIOADIAK-TGDLWEHR-RQ 117
    || : : : : : |||| : |||| : |||| : |||| : |||| : |||| :
Db 57 NVKDRIVAMTSTGVLRGH-EGSPVIGDVMVYHPPFNRFVAFALDLNDNGKILMYREPOQ 115

QY 118 LPNI 121
    || :
Db 116 DPNV 119

```

```

RESULT 14
OUIA_XANCU STANDARD; PRT: 790 AA.
ID OUIA_XANCU
AC 09XD78:
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Probable quinate dehydrogenase [pyrroloquinoline-quinone]
DE (EC 1.1.99.25).
CN OUMA.
OS Xanthomonas campestris (pv. juglandis).
OC Bacteria; Proteobacteria; gamma subphylum; Xanthomonas group;
OX NCBI_TaxID=44291;
RN NCBI_TaxID=44291;
RP SEQUENCE FROM N.A.
RC STRAIN=C5;
RA MEDLINE=20063481; PubMed=10594704;
RA Lee Y.-A., Lo Y.-C., Yu P.-P.;
RT "A gene involved in quinate metabolism is specific to one DNA homology
RT group of Xanthomonas campestris."
RL J. Appl. Microbiol. 87:649-658(1999).
CC -1- CATALYTIC ACTIVITY: Quinate + pyrroloquinoline-quinone = 5-
CC dehydroquinone + reduced pyrroloquinoline-quinone.
CC -1- COFACTOR: PQQ (BY SIMILARITY).
CC -1- PATHWAY: QUINIC ACID CATABOLIC PATHWAY; FIRST STEP. THIS PATHWAY
CC ALLOWS GROWTH OF BACTERIA WITH QUINATE BY ITS CONVERSION TO
CC PROTOCATECHUATE AND SUBSEQUENT METABOLISM BY THE BETA-KETODIPATE
CC PATHWAY.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (potential).
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL PQQ DEHYDROGENASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
CC EMBL: AF109471; AAD38453.1; -.
CC InterPro: IPR001479; Bac_PQQ.
CC InterPro: IPR002372; Bac_PQQ_repeat.
CC Pfam: PF01011; Bacterial_PQQ_6.
CC PROSITE: PS00363; BACTERIAL_PQQ_1; FALSE_NEG.
CC PROSITE: PS00364; BACTERIAL_PQQ_2; FALSE_NEG.
CC Oxidoreductase; PQQ; Quinate metabolism; Transmembrane.
FT TRANSMEM 22 42 POTENTIAL.
FT TRANSMEM 48 68 POTENTIAL.
FT TRANSMEM 77 94 POTENTIAL.
FT TRANSMEM 106 126 POTENTIAL.
SQ SEQUENCE 790 AA; 82896 MW; B75F29B52A4FE6F CRC64;

```

Query Match 13.7%; Score 90; DB 1; Length 790;
 Best Local Similarity 27.9%; Pred. No. 0.29;
 Matches 36; Conservative 18; Mismatches 53; Indels 22; Gaps 6;

```

OY 9 ASAGALLA-LAFAQVTPYDE---LLANPAG---EWISYQONENRHSPLTOI 59
   | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 110 ATYALAVVAGIGMFPVHPVAGNMGKTAVPQSGVQONMSAYGNTDGSFALDOI 169
   | : | : | : | : | : | : | : | : | : | : | : | : | : |
OY 60 TTEWVGOLQVLMARGMQPKV-----QVTPLIHDSVMTLANPGDVIQAIDAKTGD 109
   | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 170 NRSN-GRPAGSGPPTTPELANSNDGAGADQLTFLQVGEKVFICPNNLIALDASTGK 228
   | : | : | : | : | : | : | : | : | : | : | : | : | : |
OY 110 LIWEHROL 118
   | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 229 QLM--RREI 235

```

```

RESULT 15
CUD8_LOCM1 STANDARD; PRT: 148 AA.
ID CUD8_LOCM1
AC P11734:
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Cuticle protein 8 (LM-8) (LM-ACP 8).
OS Locusta migratoria (Migratory locust).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Insecta;
OC Pterygota; Neoptera; Orthoptera; Orthoptera; Caelifera;
OC Acridomorpha; Acridoidea; Acrididae; Locusta.
OX NCBI_TaxID=7004;
RN NCBI_TaxID=7004;
RP SEQUENCE.
RA MEDLINE=90073593; PubMed=2590176;
RA Klarskov K., Hoelrup P., Andersen S.O., Roepstorff P.;
RT "Plasma desorption mass spectrometry as an aid in protein sequence
RT determination. Application of the method on a cuticular protein from
RT the migratory locust (Locusta migratoria).";
RL Biochem. J. 262:923-930(1989).
RN [2]
RP SEQUENCE OF 1-56.
RA MEDLINE=86108304; PubMed=3943519;
RA Hoelrup P., Andersen S.O., Roepstorff P.;
RT "Isolation, characterization, and N-terminal sequence studies of
RT cuticular proteins from the migratory locust, Locusta migratoria.";
RL Eur. J. Biochem. 154:153-159(1986).
CC -1- FUNCTION: COMPONENT OF THE CUTICLE OF MIGRATORY LOCUST WHICH
CC CONTAINS MORE THAN 100 DIFFERENT STRUCTURAL PROTEINS.
CC -1- DOMAIN: THE TETRAPEPTIDE (A-A-P-[AV]) REPEATS FOUND THROUGHOUT THE
CC PROTEIN ARE ALSO PRESENT IN MANY PROTEINS CONSTITUTING THE
CC PROTECTIVE ENVELOPE OF OTHER SPECIES.
CC -1- SIMILARITY: BELONGS TO THE CUTICLE CONSENSUS DOMAIN.
CC
CC PIR: B24802; B24802.
CC PIR: S05638; S05638.
CC InterPro: IPR000618; Insect_cuticle.
CC Pfam: PF00379; Insect_cuticle.
CC PROSITE: PS00947; CUTICLE.
CC PROSITE: PS00233; CUTICLE.
CC structural protein; Cuticle; Repeat.
FT REPEAT 16 19 1.
FT REPEAT 22 22 2.
FT REPEAT 28 31 3.
FT REPEAT 37 40 4.
FT REPEAT 44 47 5.
SQ SEQUENCE 148 AA; 15224 MW; C6EADC27C593ACE6 CRC64;

```

Query Match 11.9%; Score 78.5; DB 1; Length 148;
 Best Local Similarity 23.1%; Pred. No. 0.61;
 Matches 24; Conservative 17; Mismatches 20; Indels 43; Gaps 3;

```

OY 8 WASAGALLA-LAFAQVTPYDELLANPAGEWISYQONENRHSPLTOITTEWVQL 67
   | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 27 VAAVAKAVVAPVAVVAPVAKAVVAP-----VAYPKYFENG 66
   | : | : | : | : | : | : | : | : | : | : | : | : | : |
OY 68 QLVARGMQPKVQVTPLIHDSVMTLANPGDVIQAIDAKTGLI 111
   | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 67 -----VHD-----AHGDIKQSGEARDGDV 87

```

Search completed: May 24, 2002, 10:15:48
 Job time: 766 sec

Fri May 24 11:27:14 2002

us-08-934-506a-5_copy_1_125.rsp

```

QY 10 SAGLALLIAAFPO-----VPEVTDELLANPAGEISYSGQEN 50
    | : | : | | | | | | | | | | | | | | | | | | | |
Db 16 STALLIASIGFAOHANNAAPESKAGQSAIENFOPYADLGLAKPPANPILIRGTQG 75
    | : | : | | | | | | | | | | | | | | | | | | | |
QY 51 YRHSPIQITTEENVGQLOLVARGMOPGVQVPTLIDGVYLANPGSDVIAQIDAKTGDL 110
    : : | | | | | | | | | | | | | | | | | | | | | |
Db 76 WGYSPLODKINDVNGDLOLVMSRMEPQSGNEGAIAIYNGVIFLGNTNDVIAQIDAKTGSL 135
    | : | : | | | | | | | | | | | | | | | | | | | |
QY 111 IWEHRROLPPN 120
    | | | : | | : | | |
Db 136 IWEYTRRKLLPS 145
    | | | : | | : | | |

```

RESULT 2

Q53362 ID Q53362 PRELIMINARY; PRT; 742 AA.

AC Q53362; Q44159;

DT 01-NOV-1996 (TREMBLrel. 01, Created)

DT 01-NOV-1996 (TREMBLrel. 01, last sequence update)

DT 01-OCT-2001 (TREMBLrel. 18, last annotation update)

DE DEHYDROGENASE SUBUNIT OF ALCOHOL DEHYDROGENASE.

OS Acetobacter pasteurianus (Acetobacter turbidans).

OC Bacteria; Proteobacteria; alpha subdivision; Acetobacteraceae;

OC Acetobacter.

OX NCBI_TaxID=438;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=NC11380;

RA MEDLINE=94042848; PubMed=822628;

RA Takemura H., Kondo K., Horinouchi S., Beppu T.;

RT "Induction by ethanol of alcohol dehydrogenase activity in Acetobacter pasteurianus";

RL J. Bacteriol. 175:6857-6866(1993).

DR EMBL; D13893; BAA40252.1; -.

DR HSSP; Q924J7; 1FLG.

DR InterPro: IPR001479; Bac_PQQ.

DR InterPro: IPR002372; Bac_PQQ_repeat.

RT Pfam: PF01011; Bacterial_PQQ; 6.

DR PROSITE: PS00363; BACTERIAL_PQQ_1; 1.

DR PROSITE: PS00364; BACTERIAL_PQQ_2; 1.

SQ SEQUENCE 742 AA; 81680 MW; 031C4E9B813EC98 CRC64;

Query Match 28.9%; Score 190; DB 2; Length 742;

Best Local Similarity 38.8%; Pred. No. 1.8e-10;

Matches 45; Conservative 18; Mismatches 49; Indels 4; Gaps 3;

QY 11 AGALALLAPAFPAQVPTVDL--ANPAGEMISYGNQENYRHSPLQITTEVNGOLO 68

DB 23 AALPAPYAPARADGGNGTGEALIHADHPENWLSIGRTYSQKRSPLDQIRSNVGDJK 82

QY 69 LVWARGMPGKVO-VTPLIHDGVWYLANPGDIVIOAIDAKTGLIWEHRRLP-NIA 122

DB 83 LAMVYLTDTNRGEATPLVVDGIMVATWMSKMEALDADATGKLLWQYDPKVGNTA 138

RESULT 3

Q9AF95 ID Q9AF95 PRELIMINARY; PRT; 691 AA.

AC Q9AF95;

DT 01-JUN-2001 (TREMBLrel. 17, Created)

DT 01-DEC-2001 (TREMBLrel. 19, last sequence update)

DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)

DE 1-BUTANOL DEHYDROGENASE BDH.

OS Pseudomonas butanovora.

OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;

OC Pseudomonas.

OX NCBI_TaxID=86174;

RN [1]

RP SEQUENCE FROM N.A.

RA Vangnai A.S., Arp D.J., Sayavedra-Soto L.A.;

RT "Characterization of the expression of two distinct alcohol dehydrogenases involved in butane metabolism in Pseudomonas butanovora";

RT Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.

RL EMBL; AF355798; AAK27220.2; -.

SQ SEQUENCE 691 AA; 75070 MW; 4FC7FD20CDA14E64 CRC64;

Query Match 27.2%; Score 178.5; DB 2; Length 691;

Best Local Similarity 37.8%; Pred. No. 2.3e-09;

Matches 31; Conservative 22; Mismatches 28; Indels 1; Gaps 1;

QY 39 GEMISYGNQENYRHSPLQITTEVNGOLOLVWARGMPGK-VQVTPLIHDGVWYLANPG 97

DB 41 GEMRTHGDDAGTRISPLAQITPDAKELGLVWSTDESSRGVEATPIVDGVMTVTAFW 100

QY 98 DVIOAIDAKTGLIWEHRRLP 119

DB 101 SVVHALDVRSGRKRLWTPDEVP 122

RESULT 4

Q9AGW3 ID Q9AGW3 PRELIMINARY; PRT; 623 AA.

AC Q9AGW3;

DT 01-JUN-2001 (TREMBLrel. 17, Created)

DT 01-JUN-2001 (TREMBLrel. 17, last sequence update)

DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)

DE 1-BUTANOL DEHYDROGENASE BOH PRECURSOR.

OS Pseudomonas butanovora.

OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;

OC Pseudomonas.

OX NCBI_TaxID=86174;

RN [1]

RP SEQUENCE FROM N.A.

RA Sayavedra-Soto L.A., Vangnai A.S., Arp D.J.;

RT "Characterization of the expression of two distinct alcohol dehydrogenases involved in butane metabolism in Pseudomonas butanovora";

RT Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.

RL EMBL; AF326086; AAK15506.1; -.

DR HSSP; Q924J7; 1FLG.

DR InterPro: IPR002372; Bac_PQQ_repeat.

DR Pfam: PF01011; Bacterial_PQQ; 4.

KW Signal.

FT SIGNAL.

SQ SEQUENCE 623 AA; 67553 MW; A573A1C64AC12D55 CRC64;

Query Match 26.5%; Score 174; DB 2; Length 623;

Best Local Similarity 38.0%; Pred. No. 5.6e-09;

Matches 49; Conservative 18; Mismatches 46; Indels 16; Gaps 5;

QY 2 KPTSLNAGALALLAPAFPAQVPTVDL--ANPAGEMISYGNQENYRHSPLQIT 59

DB 7 KPFALRAIVATAALSLPA-AAVTDVTDWEDIANDKTTGDDVLTGGLKAQGHSPKAI 65

QY 60 TTEVNGOLOLVMA-----RCMQPGKVOVTPLIHDGVWYLANPGDIVIOAIDAKTGLI 111

DB 66 NTDVNVANLVPAWSFSGEKORGE--AQV--LVHDGVYATASYSRIFAIDARSGKRL 120

QY 112 WEHRRLP 120

DB 121 WEYNARLPD 129

RESULT 5

Q9KH03 ID Q9KH03 PRELIMINARY; PRT; 698 AA.

AC Q9KH03;

DT 01-OCT-2000 (TREMBLrel. 15, Created)

DT 01-DEC-2000 (TREMBLrel. 15, last sequence update)

DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)

DE TETRAHYDROFURFURYL ALCOHOL DEHYDROGENASE.

OS Alcaligenes eutrophus (Ralstonia eutropha).

OC Bacteria; Proteobacteria; beta subdivision; Ralstonia group;

OC Ralstonia.

OX NCBI_TaxID=510;

RN [1]

RP SEQUENCE FROM N.A.

RA MEDLINE=21125557; PubMed=11222593;

RA Zarr G., Schrader T., Andreessen J.R.;

RT "Catalytic and Molecular Properties of the Quinohemoprotein Tetrahydrofurfuryl Alcohol Dehydrogenase from Ralstonia eutropha Strain Bo.";

RT J. Bacteriol. 183:1954-1960(2001).

RL EMBL; AF277373; AAF86335.1; -.

DR HSSP; Q924J7; 1FLG.

CC -1- SUBUNIT: MONOMER.
 CC -1- SUBCELLULAR LOCATION: PERIPLASMIC (POTENTIAL).
 CC -1- INDUCTION: BY ETHANOL AND BUTANOL.
 CC -1- SIMILARITY: TO ALCOHOL DEHYDROGENASE (ACCEPTOR) PROTEINS.
 DR EMBL: X61889; CA5/464.1; .
 DR HSP: Q924J7; 1FLG.
 DR InterPro: IPR002372; Bac_PQQ_repeat.
 DR InterPro: IPR003088; Cyt_C1.
 DR InterPro: IPR002329; Cyt_C1C.
 DR Pfam: PF01011; Bacterial_PQQ; 6.
 DR Pfam: PF00034; cytochrome_c; 1.
 DR PRINTS: PR00605; CYTOCHROME_C1.
 DR Signal: PQQ; Heme; Calcium; Oxidoreductase; Periplasmic.
 FT SIGNAL 1 31
 FT CHAIN 32 708
 FT TYPE I.
 FT BINDING 635 635 HEME (COVALENT) (BY SIMILARITY).
 FT BINDING 638 638 HEME (COVALENT) (BY SIMILARITY).
 FT METAL 639 639 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
 SQ SEQUENCE 708 AA; 76822 MW; 99AB54BDD6ACCAB3 CRC64;.

Query Match 25.2%; Score 165.5; DB 2; Length 708;
 Best Local Similarity 31.8%; Pred. No. 4.6e-08;
 Matches 41; Conservative 23; Mismatches 50; Indels 15; Gaps 4;

OY 3 PSLIMASAGALALAPAFAPVPTD-----DELLAMP-AGENISYQONEN 50
 DB 12 PGRWVWLLAACIG--SAAAFQGTGPAQAQAAAVRVDCDFRANAARTPTDPTGVYAE 69
 OY 51 YRHSPLQITTEENVGQQLVWARGMQPK-VQVPLIHGVYLANPGDYQAIIDAKTGD 109
 DB 70 TRYSLLDINANVKDGLAWSYNLESTRGEATPVVDGIMYASMSVYHAIDRTGN 129
 OY 110 LIWEHRRQL 118
 DB 130 RIMTYDPOI 138

RESULT 8
 OY9F9U2 PRELIMINARY; PRT; 695 AA.
 AC Q9F9U2;
 DT 01-MAR-2001 (Tremblrel. 16, Created)
 DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
 DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
 DE ALCOHOL DEHYDROGENASE.
 GN ADH.
 OS Pseudomonas stutzeri (Pseudomonas perfectomarina).
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 OC Pseudomonas.
 OX NCBI_TaxID=316;
 RN [1]
 RC SEQUENCE FROM N.A.
 RA STRAIN=BC-2;
 RA Chang C.-H., Herrick J.B., Okinaka R.T., Brainard J.B.,
 RA Terwilliger T.C.;
 RT Identification and characterization of genes activated by 2-
 RT chloroethanol in Pseudomonas stutzeri BC-2.;
 RL Submitted (Aug-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF176640; AAG09249.1; .
 DR HSP: Q924J7; 1FLG.
 DR InterPro: IPR002372; Bac_PQQ_repeat.
 DR InterPro: IPR000345; CytC_heme_bind.
 DR InterPro: IPR003088; Cyt_C1.
 DR Pfam: PF01011; Bacterial_PQQ; 6.
 DR Pfam: PF00034; cytochrome_c; 1.
 DR PROSITE: PS00190; CYTOCHROME_C; UNKNOWN.1.
 SQ SEQUENCE 695 AA; 75842 MW; D29A698A642E1A83 CRC64;

Query Match 23.7%; Score 156; DB 2; Length 695;
 Best Local Similarity 33.6%; Pred. No. 4e-07;

Matches 36; Conservative 22; Mismatches 45; Indels 4; Gaps 3;
 OY 15 ALLAPAPAFQVPTDELL--ANPAGEMISYQONENYRSPITQITTEVWGLQVMA 72
 DB 16 ALIVA-AGAQAQAKVDEAIFRASEDGESEMLSHGRITAYAGRRSPKQIDAGVNGKGLAWY 74
 OY 73 ROMOPKQ-VQVPLIHGVYLANPGDYQAIIDAKTGDLIWEHRRQL 118
 DB 75 LDLENKRGLEATPLVSDGLVYASLSWSRVMAVDLRSGRILMQDFDPV 121

RESULT 9
 OY98KF6 PRELIMINARY; PRT; 785 AA.
 AC Q98KF6;
 DT 01-OCT-2001 (Tremblrel. 18, Created)
 DT 01-OCT-2001 (Tremblrel. 18, Last sequence update)
 DT 01-OCT-2001 (Tremblrel. 18, Last annotation update)
 DE GLUCOSE DEHYDROGENASE.
 GN MLI1500.
 OS Rhizobium loti (Mesorhizobium loti).
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Rhizobiaceae; Mesorhizobium.
 OX NCBI_TaxID=381;
 RN [1]
 RC SEQUENCE FROM N.A.
 RA STRAIN=MAFE303099;
 RX MEDLINE=21082930; PubMed=11214968;
 RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
 RA Watanabe A., Ideasa K., Ishikawa A., Kawashima K., Kimura T.,
 RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
 RA Mochizuki Y., Nakayama S., Nakazaki N., Shimo S., Sugimoto M.,
 RA Takeuchi C., Yamada M., Tabata S.;
 RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
 RT Mesorhizobium loti".
 RL DNA Res. 7:331-338(2000).
 DR EMBL: AP002997; BAB48838.1; .
 DR InterPro: IPR002372; Bac_PQQ_repeat.
 DR Pfam: PF01011; Bacterial_PQQ; 7.
 KW Complete proteome.
 SQ SEQUENCE 785 AA; 84083 MW; IDC224B7A1C25470 CRC64;

Query Match 23.6%; Score 155; DB 16; Length 785;
 Best Local Similarity 31.7%; Pred. No. 5.8e-07;
 Matches 46; Conservative 13; Mismatches 46; Indels 40; Gaps 4;
 OY 8 WASAG-----ALALAPAFAPVPTD-----ELLANPAGE 40
 DB 111 WARRGLAGPDRAPLILAVLSAVAGYSMTADKGIIDALDVKIPKANLGNDVPAGE 170
 OY 41 WISYGQONENYRSPITQITTEENVGQQLVWARGMQPKV-----QVPLIHG 89
 DB 171 WHYGRTOFGQRYSPDLQITPDVANLQPAWT--YRTGVKGGDDIGETTYVTPKIGD 228
 OY 90 VMYLANPGDYQAIIDAKTGDLIWEH 114
 DB 229 TLVICTHNFALAVDAATGKEKMWY 253

RESULT 10
 OY92RB3 PRELIMINARY; PRT; 777 AA.
 AC Q92RB3;
 DT 01-DEC-2001 (Tremblrel. 19, Created)
 DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE PROBABLE GLUCOSE DEHYDROGENASE (PYRROLOQUINOLINE-QUINONE) PROTEIN
 DE (EC 1.1.99.17).
 OS Rhizobium meliloti (Sinorhizobium meliloti).
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Rhizobiaceae; Sinorhizobium.
 OX NCBI_TaxID=382;

[illegible][illegible][illegible]

```

Query Match Similarity 21.7% 17: Missed
Best Local 43: Conservative 11
Matches
6 LHM-----ASAGALIALIAFAFQVPTPYDELINP-----
111
0Y

```

1 125. ispt

161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000

[illegible][illegible]

RESULT 13 PRELIMINARY: Created) (sequence update)
O92NW9 092NW9 19 last seqequence updates) (EC
ID O92W9 19 last seqequence updates) (EC
AC O92W9-2001 (TEMPRELrel. 19, last amonction updates)
DT 01-DEC-2001 (TEMPRELrel. 19, last amonction updates)
PT 01-DEC-2001 (TEMPRELrel. 19, last amonction updates)
DE PUVATIVE METHANOL DEHYDROGENASE PROTEIN, LARG SUBUNIT (EC
DT FUVATIVE METHANOL DEHYDROGENASE PROTEIN, LARG SUBUNIT (EC
DE 1.99.8).
GN SMB20173. melliocti (Sinothizobium melliocti). Rhizobiaceae group;
GN Rhizobium psymb (megaplasmid 2). alpha subdivision; Rhizobiaceae group;
OS plasmid Proteobacteria: Sinothizobium.
OC Bacterifera: Proteobacteria: Sinothizobium.
CC Bacterifera: Proteobacteria: Sinothizobium.
OX NCBI-taxid:382;
RN NCBI-taxid:382;
RP SCQUENCE FROM N.A. Chain P., Gouzy J.,
RC STRAIN-1021. Pubmed-1181431;
RX MEDIAN-21396508. Pubmed-1181431;
RA MEDIAN T.M., Weidner S., Wong K., Buinmester J., Couste A.,
RA FURHOELTER F., Puellier A., Hernandez-Lucas I., Becker A.,
RA Goiding B., Puellier A., Of the 1,683-kb psymb megaplasmid from the N2
RT "the complete sequence of the 1,683-kb psymb megaplasmid from the N2

Fri May 24 11:27:15 2002

RT fixing endosymbiont Sinorhizobium meliloti.
 NL Proc. Natl. Acad. Sci. U.S.A. 98:9889-9894(2001).
 KR EMBL: U603642; CAC48573.1; A. 98:9889-9894(2001).
 SO OXIDOREDUCTASE; PLASMID; Hypothetical protein; Complete proteome.

Query Match

Best Local Similarity 20.58; Score 134.5; DB 16; Length 601;
 Matches 38; Conservative 27; Mismatches 41; Indels 9; Gaps 7;

OY 14 LALLAPAFAYPTDEL--LAMPAGEMISYGOENYRHSPLTITTEWVGOLOLW 71
 DB 8 LAINISGGAQVAFANDELOLIDP-NOMAIQNGDIANLRYSKLDINKDNVGLQVAM 66
 OY 72 --ARGMQPKVQVPTLIDHCVMTLAMP-GDVIOAID-ATGDLWEHR-NOLPNI 121
 DB 67 TFEIVLNGH-EGSPVYDILMYVHTFPVYVALDLSKDOIVWKYEKP-NODPNV 120
 RESULT 14
 ID 09X255
 AC 09X255
 DT 01-NOV-1999 (TREMBLrel. 12, Created)
 DE 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
 DE GLUCOSE DEHYDROGENASE.
 GN GDB.
 OS Pantoea citrea.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OX NCI
 RX NCI
 RA MEDLINE-9936331; PubMed-10376838;
 RT "gdb", a g...
 RL Pantoea citrea, encoding a second glutoprotein glucose dehydrogenase in
 DR EMBL: AF050503; AD23735.1; (1999).
 DR InterPro: IPR001479; Bac_PDO.
 DR Pfam: PF01011; Bacterial_PDO_repeat.
 DR PROSITE: PS00364; Bacterial_PDO_repeat.
 SO SEQUENCE 790 AA; 85681 MW; 08B87625E0A5F089 CRC64;

Query Match

Best Local Similarity 18.48; Score 121; DB 2; Length 790;
 Matches 31; Conservative 13; Mismatches 39; Indels 12; Gaps 3;

OY 27 PVTELLANPRAGEMISYGOENYRHSPLTITTEWVGOLOLW 112
 DB 152 PATQPVSAK--DWTAYGGTQAGDYSLSQINESVKNLQVAMTYQSGDVKRPDDGE 208
 OY 80 V--VTPLIHGVVYLANPGDVIOAIDAKTGDLIW 112
 DB 209 ITNESTLKNVNLVCTAHQIIVAIIDAKYGEKW 243
 RESULT 15
 ID 0934G0
 AC 0934G0
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DE 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DE LUPANINE HYDROXYLASE.
 GN LUP.
 OS Pseudomonas sp. DH2001
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;

us-08-934-506a-5-copy_1_125.rspt

OC Pseudomonas.
 RN NCI
 RP SEQUENCE FROM N.A.
 RA Hopper D.J.; Kaderbhai M.A.; Little A.R.; Marriott S.A.; Young M.,
 RT "Clon...
 RL hydroxylase, sequencing and analysis of the gene for lupanine
 DR EMBL: A7318095; CAC67410.1; to the EMBL/GenBank/DBJ databases.
 DR Signal.
 FT SIGNAL
 FT CHAIN
 SO SEQUENCE 27 1 26 695
 ID 0934G0
 AC 0934G0
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DE 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DE LUPANINE HYDROXYLASE.
 GN LUP.
 OS Pseudomonas sp. DH2001
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;

Query Match

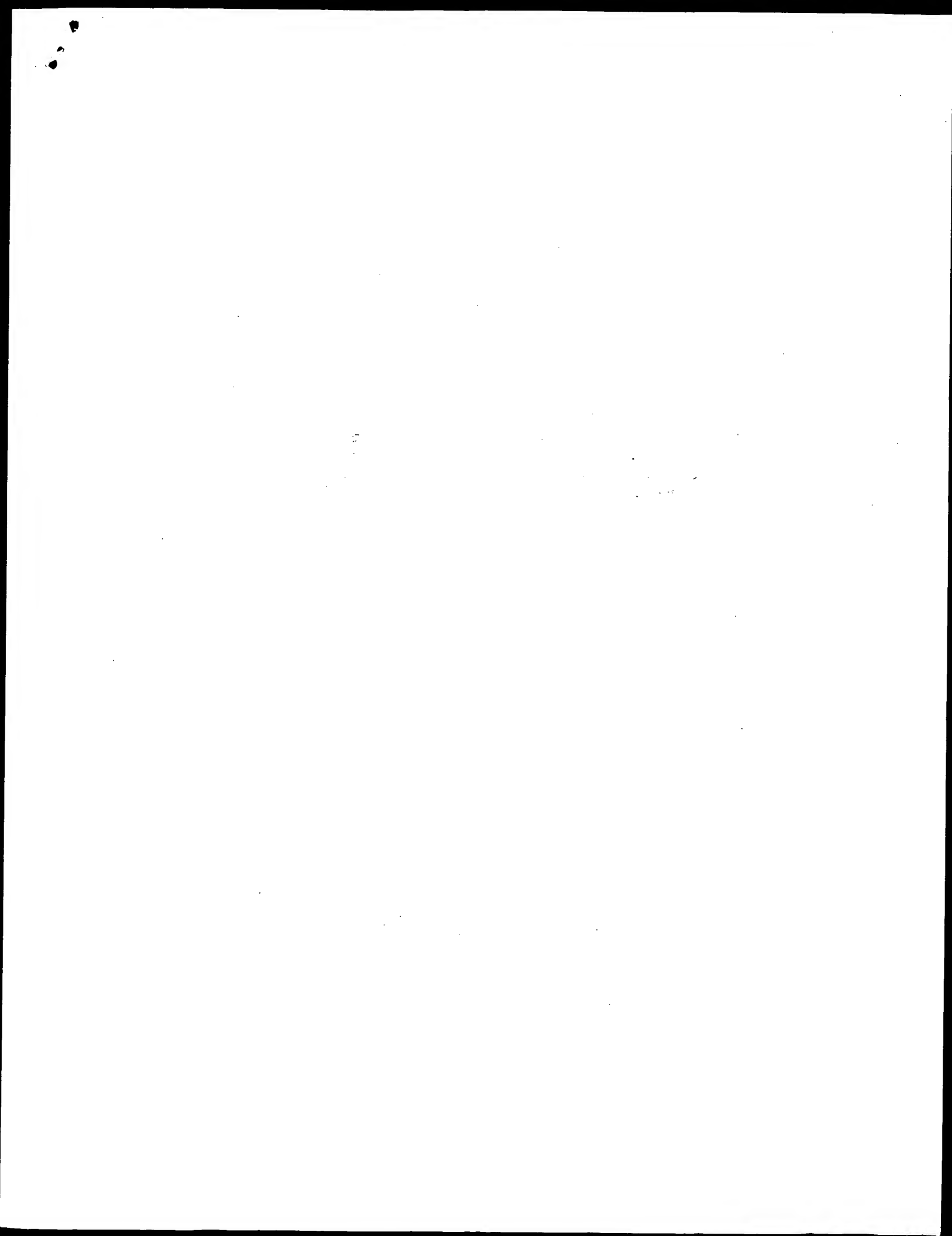
Best Local Similarity 18.08; Score 118.5; DB 2; Length 695;
 Matches 28; Conservative 8; Mismatches 41; Indels 1; Gaps 1;

OY 38 AGEWISYGOENYRHSPLTITTEWVGOLOLWARGMQPKVQV-TPLIHDSVMTLAMP 96
 DB 36 SGWMSLIGGNEOHYSALSALDKVNSVKNLGLSMTDWEADGLVGNPLVADGVITYGCGP 95
 OY 97 GDVIOAIDAKTGDLWEH 114
 DB 96 PKIYANDLKTIRKINLMTY 113

Search completed: May 24, 2002, 10:17:30
 Job time: 743 sec

Fri May 24 11:27:15 2002

us-08-934-506a-5_copy_1_125.rspt



GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2002, 10:05:00 ; Search time 117.28 Seconds

(without alignments)
121.227 Million cell updates/sec

Title: US-08-934-506a-5_COPY_1_128

Perfect score: 673
Sequence: 1 MKPSTLWMSAGALALLAAP.....DLWEHRRQLPNATLNSFG 128

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_032802:*

- 1: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1980.DAT:*
- 2: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1981.DAT:*
- 3: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1982.DAT:*
- 4: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1983.DAT:*
- 5: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1984.DAT:*
- 6: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1985.DAT:*
- 7: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1986.DAT:*
- 8: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1987.DAT:*
- 9: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1988.DAT:*
- 10: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1989.DAT:*
- 11: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1990.DAT:*
- 12: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1991.DAT:*
- 13: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1992.DAT:*
- 14: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1993.DAT:*
- 15: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1994.DAT:*
- 16: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1995.DAT:*
- 17: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1996.DAT:*
- 18: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1997.DAT:*
- 19: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1998.DAT:*
- 20: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1999.DAT:*
- 21: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA2000.DAT:*
- 22: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	673	100.0	579	19	AAW37873
2	603	89.6	579	19	AAW37874
3	552	82.0	579	19	AAW37876
4	550.5	81.8	578	19	AAW37875
5	189.5	28.2	754	21	AAW35987
6	183	27.2	738	12	AAW35987
7	183	27.2	738	12	AAW35987
8	177	26.3	742	11	AAW20192
9	155	23.0	443	22	AAW05235
10	145.5	21.6	740	20	ABG24430
11	129	19.2	948	22	ABG25909

12	129	19.2	1510	22	ABG21573
13	76	11.3	824	18	AAW09614
14	73.5	10.9	1229	22	ABB71736
15	72.5	10.8	343	19	AAW35001
16	72.5	10.8	462	17	AAW05395
17	72	10.7	282	22	AAW1927
18	71	10.5	526	22	AAW9895
19	70	10.4	1016	22	AAU34349
20	69.5	10.3	275	21	AAW4781
21	69.5	10.3	1029	22	ABG26194
22	69.5	10.3	1029	22	ABG26568
23	69.5	10.3	1419	22	ABB59208
24	69	10.3	711	22	ABB69743
25	69	10.3	957	22	ABB62401
26	68.5	10.2	353	22	AAW47447
27	68.5	10.2	384	22	AAW85255
28	68.5	10.2	551	22	AAW85251
29	68.5	10.2	551	22	AAW85252
30	68.5	10.2	843	20	AAV34485
31	68.5	10.2	849	20	AAV34360
32	68	10.1	159	19	AAW63034
33	68	10.1	159	21	AAW14325
34	68	10.1	294	21	AAW35811
35	68	10.1	294	21	AAW15001
36	68	10.1	424	22	ABB59716
37	67.5	10.0	171	22	AAU32330
38	67.5	10.0	275	21	AAW16656
39	67.5	10.0	735	22	ABG24896
40	67.5	10.0	974	19	AAW40079
41	67.5	10.0	1041	22	ABG21594
42	67.5	10.0	1224	18	AAW26735
43	67	10.0	668	22	ABB66214
44	67	10.0	673	22	ABB5674
45	67	10.0	1407	22	ABG28517

ALIGNMENTS

RESULT 1	AAW37873	standard; Protein; 579 AA.
ID	AAW37873;	
AC	AAW37873;	
XX		
DT	10-AUG-1998	(first entry)
XX		
DE	Alcohol and/or aldehyde dehydrogenase A amino acid sequence.	
XX		
KW	Alcohol/aldehyde dehydrogenase A enzyme; recombinant organism;	
KW	aldehyde; ketone; carboxylic acid; L-sorbose; D-sorbitol;	
KW	2-keto-L-gulononic acid; L-ascorbic; inhibition.	
OS	Gluconobacter oxydans.	
XX		
XX	Key	Location/Qualifiers
FT	Peptide	1..23
FT	Protein	/note="signal peptide"
FT		24..579
FT		/note="mature protein"
XX		
PN	EP832974-A2.	
XX		
PD	01-APR-1998.	
XX		
PF	11-SEP-1997;	97EP-0115801.
XX		
PR	19-SEP-1996;	96EP-0115001.
XX		
PA	(HOFF) HOFFMANN LA ROCHE & CO AG F.	
XX		
PI	Asakura A, Hoshino T, Ojima S, Shinjoh M, Tomiyama N;	
XX		

Novel human diago
pathogenic Staphyl
Drosophila melanog
Bacillus thermoloe
Human SH3P17 prote
Acremonium cellulo
Xanthomonas sp. ch
Staphylococcus aur
Arabidopsis thalia
Novel human diago
Novel human diago
Drosophila melanog
Drosophila melanog
MOMP p5. Haemophi
Plant porphobilinog
Plant thiorodoxin-
Plant thiorodoxin-
Porphyromonas ging
Porphyromonas ging
Mycobacterium tube
Mycobacterium tube
Chitinase, amino ac
Chitinase protein
Drosophila melanog
Novel human secret
Arabidopsis thalia
Novel human diago
Solanum tuberosum
Novel human diago
Novel human diago
Staphylococcus car
Drosophila melanog
Drosophila melanog
Novel human diago

DR WPI: 1998-195228/18.
 DR N-PSDB: AAV29051.
 XX
 PT Recombinant Gluconobacter oxydans alcohol and/or aldehyde
 PT dehydrogenase enzyme(s) - useful for converting substrate(s), e.g.
 PT L-sorbose or D-sorbitol to 2-keto-L-gulonic acid
 XX
 PS Claim 1; Pages 35-37; 59pp; English.
 XX
 CC This is the amino acid sequence for the Gluconobacter oxydans
 CC alcohol and/or aldehyde dehydrogenase A enzyme. The enzymes
 CC or recombinant organisms can be used to convert suitable substrates
 CC to aldehydes, ketones or carboxylic acids, especially to convert
 CC L-sorbose or D-sorbitol to 2-keto-L-gulonic acid, which can be
 CC converted to L-ascorbic acid by standard procedures. The derivatives
 CC of ADH enzymes have desired substrate specificity, higher affinity
 CC to a substrate, lower affinity to an inhibitory compound, higher
 CC stability against temperature and/or pH and higher catalytic speed.
 CC
 SQ Sequence 579 AA;

Query Match 100.0%; Score 673; DB 19; Length 579;
 Best Local Similarity 100.0%; Pred. No. 8.3e-65;
 Matches 128; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MKPTSLMASGALALLAFAFAQVTPVDELLANPPAGEWISYGNQENYRHSPLTQIT 60
 DB 1 mptslmasagalallaafafaqvtprtdellanppagewisynqenyrhsplqt 60
 QY 61 TENVCGLQLVWARGMQPKVQVTPPLIHGVMYLANPGDVYQAIDAKTGLIWEHRQLPN 120
 DB 61 tenvcglqlvwargmqpkvqvtpplihgvmylanpgdviqaidaktgdlwehrrqlpn 120
 QY 121 IATLNSFG 128
 DB 121 iatlnsfq 128

RESULT 2
 AAM37874
 ID AAM37874 standard; Protein: 579 AA.
 AC AAM37874;
 DT 10-AUG-1998 (first entry)
 DE Alcohol and/or aldehyde dehydrogenase A' amino acid sequence.
 XX
 KW Alcohol/aldehyde dehydrogenase A' enzyme; recombinant organism;
 KW aldehyde; ketone; carboxylic acid; L-sorbose; D-sorbitol;
 KW 2-keto-L-gulonic acid; L-ascorbic; inhibition.
 OS Gluconobacter oxydans.
 FH Key
 FT Peptide 1..23
 FT /note= "signal peptide"
 FT Protein 24..579
 FT /note= "mature protein"
 XX
 EN EP832974-A2.
 PD 01-APR-1998.
 PF 11-SEP-1997; 97EP-0115801.
 PR 19-SEP-1996; 96EP-0115001.
 PA (HOFF) HOFFMANN LA ROCHE & CO AG F.
 PI Asakura A, Hoshino T, Ojima S, Shinjoh M, Tomiyama N;
 XX

DR WPI: 1998-195228/18.
 DR N-PSDB: AAV29052.
 XX
 PT Recombinant Gluconobacter oxydans alcohol and/or aldehyde
 PT dehydrogenase enzyme(s) - useful for converting substrate(s), e.g.
 PT L-sorbose or D-sorbitol to 2-keto-L-gulonic acid
 XX
 PS Claim 1; Pages 38-40; 59pp; English.
 XX
 CC This is the amino acid sequence for the Gluconobacter oxydans
 CC alcohol and/or aldehyde dehydrogenase A' enzyme. The enzymes
 CC or recombinant organisms can be used to convert suitable substrates
 CC to aldehydes, ketones or carboxylic acids, especially to convert
 CC L-sorbose or D-sorbitol to 2-keto-L-gulonic acid, which can be
 CC converted to L-ascorbic acid by standard procedures. The derivatives
 CC of ADH enzymes have desired substrate specificity, higher affinity
 CC to a substrate, lower affinity to an inhibitory compound, higher
 CC stability against temperature and/or pH and higher catalytic speed.
 CC
 SQ Sequence 579 AA;

Query Match 89.6%; Score 603; DB 19; Length 579;
 Best Local Similarity 91.4%; Pred. No. 4.7e-58;
 Matches 117; Conservative 3; Mismatches 8; Indels 0; Gaps 0;
 QY 1 MKPTSLMASGALALLAFAFAQVTPVDELLANPPAGEWISYGNQENYRHSPLTQIT 60
 DB 1 mktsslmasagalallaafafaqvtprtdellanppagewisynqenyrhsplqt 60
 QY 61 TENVCGLQLVWARGMQPKVQVTPPLIHGVMYLANPGDVYQAIDAKTGLIWEHRQLPN 120
 DB 61 tenvcglqlvwargmqpkvqvtpplihgvmylanpgdviqaidaktgdlwehrrqlpn 120
 QY 121 IATLNSFG 128
 DB 121 iatlnsfq 128

RESULT 3
 AAM37876
 ID AAM37876 standard; Protein: 579 AA.
 AC AAM37876;
 DT 10-AUG-1998 (first entry)
 DE Alcohol and/or aldehyde dehydrogenase B amino acid sequence.
 XX
 KW Alcohol/aldehyde dehydrogenase B enzyme; recombinant organism;
 KW aldehyde; ketone; carboxylic acid; L-sorbose; D-sorbitol;
 KW 2-keto-L-gulonic acid; L-ascorbic; inhibition.
 OS Gluconobacter oxydans.
 FH Key
 FT Peptide 1..23
 FT /note= "signal peptide"
 FT Protein 24..579
 FT /note= "mature protein"
 XX
 EN EP832974-A2.
 PD 01-APR-1998.
 PF 11-SEP-1997; 97EP-0115801.
 PR 19-SEP-1996; 96EP-0115001.
 PA (HOFF) HOFFMANN LA ROCHE & CO AG F.
 PI Asakura A, Hoshino T, Ojima S, Shinjoh M, Tomiyama N;
 XX

DR MPI: 1998-195228/18.
DR N-PSDB: AAV29054.
XX
XX Recombinant Gluconobacter oxydans alcohol and/or aldehyde
PT dehydrogenase enzyme(s) - useful for converting substrate(s), e.g.
PT L-sorbose or D-sorbitol to 2-keto-L-gulononic acid
XX
PS Claim 1; Pages 44-46; 59pp; English.
XX
XX This is the amino acid sequence for the gluconobacter oxydans
CC alcohol and/or aldehyde dehydrogenase B enzyme. The enzymes
CC or recombinant organisms can be used to convert suitable substrates
CC to aldehydes, ketones or carboxylic acids, especially to convert
CC L-sorbose or D-sorbitol to 2-keto-L-gulononic acid, which can be
CC converted to L-ascorbic acid by standard procedures. The derivatives
CC of AADH enzymes have desired substrate specificity, higher affinity
CC to a substrate, lower affinity to an inhibitory compound, higher
XX stability against temperature and/or pH and higher catalytic speed.
SQ Sequence 579 AA;

DR WPI: 1998-195228/18.
DR N-PSDB: AAV29053.
XX
XX Recombinant Gluconobacter oxydans alcohol and/or aldehyde
PT dehydrogenase enzyme(s) - useful for converting substrate(s), e.g.
PT L-sorbose or D-sorbitol to 2-keto-L-gulononic acid
XX
XX Claim 1; Pages 41-43; 59pp; English.
PS
XX This is the amino acid sequence for the Gluconobacter oxydans
CC alcohol and/or aldehyde dehydrogenase A' enzyme. The enzymes
CC or recombinant organisms can be used to convert suitable substrates
CC to aldehydes, ketones or carboxylic acids, especially to convert
CC L-sorbose or D-sorbitol to 2-keto-L-gulononic acid, which can be
CC converted to L-ascorbic acid by standard procedures. The derivatives
CC of ADH enzymes have defined substrate specificity, higher affinity
CC to a substrate, lower affinity to an inhibitory compound, higher
XX stability against temperature and/or pH and higher catalytic speed.
XX
SQ Sequence 578 AA:

PT production of 2-keto-L-gulononic acid and L-sorbose from D-sorbitol
 XX
 PS Claim 1; Fig 8; 96pp; English.
 XX
 CC This invention relates to an isolated membrane-bound sorbitol
 CC dehydrogenase (SDH) from *Glucobacter oxydans*. The invention includes
 CC polynucleotides (AAC83153 - AAC83154) encoding the 3 subunits of SDH
 CC (AAB35987 - AAB35989). Also included in the invention are two
 CC polynucleotide sequences AAC83156 and AAC83157 which contain the subunit
 CC coding regions. The sorbitol dehydrogenase polynucleotide sequences are
 CC useful for producing L-sorbose from D-sorbitol and for increasing the
 CC production of 2-keto-L-gulononic acid by transforming a host cell,
 CC especially *Glucobacter* with the DNA and selecting the transformed host
 CC cell. Note: An additional 66 DNA sequences (fragments of the SDH subunit
 CC genes) encoding fragments of SDH are specifically claimed, however these
 CC sequences are not given separately in the specification but are included
 CC in sequences AAC83156 and AAC83157.
 CC
 XX
 SQ Sequence 754 AA:

Query Match 28.2%; Score 189.5; DB 21; Length 754;
 Best Local Similarity 39.0%; Pred. No. 3.9e-12;
 Matches 46; Conservative 17; Mismatches 46; Indels 9; Gaps 3;

OY 14 LLLAAPAFAYVTV-----TDELLANPPAGEWISYQONENYRHSPLQITTEVNGQ 66
 Db 17 lgcAaalfatcspalaeatgtaltndqbpqdmgyrlysegrypldqtkdnasn 76
 OY 67 LQLVWARGMQPKGVYV-TPLIHDGVMYLANPQDVYQAIADAKTGDILWEHRRQLP-NIA 122
 Db 77 lklawhyldltmrgegrpllydgyvmyatltwskmkaldatgkllwysdtkvpqnia 134

RESULT 6
 AAR13993 standard; Protein: 738 AA.

XX AAR13993;
 XX
 DT 09-DEC-1991 (first entry)
 XX
 DE A.altoacetigenes membrane-bound ADH 72kd sub-unit.
 XX
 KW alcohol dehydrogenase complex; carboxylic acid production.
 XX
 OS Acetobacter altoacetigenes.
 XX
 OS EP448969-A.
 XX
 PN 02-OCT-1991.
 XX
 PD 26-FEB-1991; 91EP-0102793.
 XX
 PF 26-FEB-1991; 91EP-0102793.
 XX
 PR 26-MAR-1990; 90JP-0073440.
 PR 26-FEB-1990; 90JP-0042391.
 XX
 PA (NAKA-) NAKANO VINEGAR KK.
 XX
 PI Tamaki T, Takemura H, Tayama K, Fukaya M, Okumura H;
 PI Kawamura Y;
 PI WPI: 1991-289462/40.
 DR N-PSDB; AAQ13580.
 XX
 XX
 PT Gene for membrane-bound alcohol dehydrogenase complex - optd.
 PT from Acetobacter altoacetigenes, used for prodn. of enzyme for
 PT converting alcohol to acid
 XX
 PS Disclosure; Fig 3; 36pp; English.
 CC
 CC Total DNA was prepared from A.altoacetigenes MH-24, digested with
 CC PstI and SmaI and ligated to PstI-SmaI cleaved pUC18. The ligation

CC mixture was used to transform E.coli JM109. Probes were designed
 CC based on the N-terminal amino acid sequence of the ADH complex
 CC isolated from A. altoacetigenes (see AAQ13582-Q13584). The gene
 CC encoding the 72kd ADH subunit was isolated and sequenced. The
 CC directly sequenced N-terminal region of the purified 72kd sub-unit
 CC corresponds to the sequence beginning at residue 36 of the deduced
 CC sequence. This suggests that the first 35 N-terminal amino acids
 CC form a leader peptide involved in secretion of the mature 72kd
 CC protein. The deduced amino acid sequence has 77 per cent homology
 CC with the same enzyme from A. aceti K6033. See also AAQ13581.
 CC
 XX
 SQ Sequence 738 AA:

Query Match 27.2%; Score 183; DB 12; Length 738;
 Best Local Similarity 33.3%; Pred. No. 2e-11;
 Matches 41; Conservative 23; Mismatches 49; Indels 10; Gaps 3;

OY 10 SAGALALAAPAFAYVTVDE-----LLANPPAGEWISYQONENYRHSPLQIT 61
 Db 17 tagtcaallsgyatmasaddggatgcaalhhaddhpnwmygrlysdqgrpsldqlnr 76
 OY 62 ENVGQDLVWARGMQPKGVYV-TPLIHDGVMYLANPQDVYQAIADAKTGDILWEHRRQLP- 119
 Db 77 snvgnlkrlawlyldltmrgegrpllydgyvmyatltwsmmkavdatgkllwysdtkvpq 136
 OY 120 NIA 122
 Db 137 nla 139

RESULT 7
 AAR20192 standard; Protein: 738 AA.

XX AAR20192;
 XX
 DT 16-APR-1992 (first entry)
 XX
 DE ADH complex protein (mol.wt. 72.000).
 XX
 KW Alcohol dehydrogenase; acetic acid; fermentation.
 XX
 OS Acetobacter altoacetigenes NH-24.
 XX
 OS JF0326998-A.
 XX
 PN 27-NOV-1991.
 XX
 PD 26-MAR-1990; 90JP-0073440.
 XX
 PF 26-FEB-1990; 90JP-0042301.
 PR 26-MAR-1990; 90JP-0073440.
 XX
 PA (NAKA-) NAKANO SUTEN KK.
 XX
 PI WPI: 1992-019325/03.
 DR N-PSDB; AAQ20383.
 XX
 XX
 PT Alcohol dehydrogenase complex structural gene - used in plasmid
 PT and enhancing efficiency of acetic acid fermentation for
 PT transformed acetic acid bacteria
 XX
 PS Disclosure; Fig 3(1-3); 21pp; Japanese.
 CC
 CC Acetobacter transformed with the sequence encoding this protein can
 CC enhance the efficiency of acetic acid fermentation. The ADH complex
 CC can be easily extracted from the bacteria and purified and it can be
 CC used for the determination of an alcohol.
 CC See also AAQ20383-84, and -86-88.
 CC
 XX
 SQ Sequence 738 AA:

Db 168 ltwagfndpqlngltlsadatpaea-ispvadq-----dwpaygrnqeggrfispkq 218
 QY 59 ITFENYGOLOLWVA-----RGMQGRKY--QVTPLIHGVWYLANPGVIOATDAKTGD 109
 Db 219 lnadrvhnlkeavvrtgtgkqndpgeitnevcplkvgdtylctahqrlfaldasgk 278
 QY 110 LIWEHRRQLPNIATLNSF 127
 Db 279 ekwhyd---pelknesf 293

RESULT 10

AAW95019
 ID AAW95019 standard; Protein; 740 AA.

AC AAW95019;

DT 21-MAY-1999 (first entry)

DE Sorbitol dehydrogenase (SLDH) protein sequence.

KW Sorbitol dehydrogenase; SLDH; open reading frame; ORF2; L-sorbose;
 KM D-sorbitol; Vitamin C; enzyme.

OS Gluconobacter suboxydans.

FT Key Location/Qualifiers

FT Peptide 1..24

FT Protein /note= "signal sequence"
 25..740

PN EEP97984-A2.

PD 24-FEB-1999.

PF 13-AUG-1998; 98EP-0115231.

PR 21-AUG-1997; 97EP-0114432.

PA (HOFF) HOFFMANN LA ROCHE & CO AG F.

PI Hosino T, Miyazaki T, Ojima S, Shinjoh M, Tomiyama N;

DR WPI; 1999-134646/12.

DR N-PSDB; AAX21501.

PT New D-sorbitol dehydrogenase gene and recombinant protein - useful
 for production of L-sorbose, an intermediate in vitamin C production

PS Claim 1; Fig 3A-D; 39pp; English.

CC This represents a sorbitol dehydrogenase (SLDH) protein. The DNA
 CC encoding the SLDH enzyme also encodes an open reading frame (ORF2)
 CC product upstream of the SLDH open reading frame, needed for SLDH activity
 CC in vivo. Host cells transformed by a vector comprising the SLDH DNA
 CC sequence are used for the recombinant expression of the sorbitol
 CC dehydrogenase. Recombinant D-sorbitol dehydrogenase is useful for
 CC producing L-sorbose from D-sorbitol. L-sorbose is an important
 CC intermediate in vitamin C production.

XX Sequence 740 AA;

Query Match 21.6%; Score 145.5; DB 20; Length 740;
 Best Local Similarity 34.0%; Pred. No. 2.8e-07;

Matches 34; Conservative 25; Mismatches 30; Indels 11; Gaps 5;

QY 23 AOVTPVTDLLANPPAGWISYGOENORHSPLOITTEWYGOLOL--VNAKGM--QPG 78

Db 86 sqvpmappgsanpargdwvaygriddhqrtygsplseipenasklkvalvnyhtgsyprpg 145

QY 79 KV-----QVTPVLIHGVWYLANPGDVIOATDAKTGDLIW 112

Db 146 qvnxwaetpikvgdglytcsamndlik-ldpatgqdlw 184

RESULT 11

ABG25909
 ID ABG25909 standard; Protein; 948 AA.

AC ABG25909;

DT 18-FEB-2002 (first entry)

DE Novel human diagnostic protein #25900.

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KM food supplement; medical imaging; diagnostic; genetic disorder.

OS Homo sapiens.

PN WO200175067-A2.

PD 11-OCT-2001.

PF 30-MAR-2001; 2001WO-US08631.

PR 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

PA (HYSE-) HYSEQ INC.

PI Drmanac RT, Liu C, Tang YT;

DR WPI; 2001-639362/73.

DR N-PSDB; AAS90096.

PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity

PS Claim 20; SEQ ID No 56268; 103pp; English.

CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pot_sequences.

XX Sequence 948 AA;

Query Match 19.2%; Score 129; DB 22; Length 948;

Best Local Similarity 29.9%; Pred. No. 2.6e-05;

Matches 43; Conservative 17; Mismatches 30; Indels 54; Gaps 8;

QY 6 LWMAS-----AGALLAAPAFQVTPVTDLLANPPAGWISYGOENORHSPLO 58

Db 310 ltwgafndpgelngtlisadtpaea-ispvadq-----dwpaygrngqgrfslpkq 360
 QY 59 ITTENVGOLQLVMA-----RGMQPGKV--QVTPLIHGVMTLANPGDVIALDAKTGD 109
 Db 361 lnadnhnlkeavfrtgdvkqpnpgelntevtpi-----kvqd 400
 QY 110 ---LIMEHRR-----QLPNIAATL 124
 Db 401 tpllctahgrlfaaldaemdgiaatl 424

RESULT 12

ABG21573
 ID ABG21573 standard; Protein: 1510 AA.

AC ABG21573;

DT 18-FEB-2002 (first entry)

DE Novel human diagnostic protein #21564.

KM Human; chromosome mapping; gene mapping; gene therapy; forensic;

KN food supplement; medical imaging; diagnostic; genetic disorder.

OS Homo sapiens.

PN WO200175067-A2.

PD 11-OCT-2001.

PF 30-MAR-2001; 2001WO-US08631.

PR 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

PA (HYSE-) HYSEQ INC.

PI Drmanac RT, Liu C, Tang YT;

DR WPI; 2001-639362/73.

DR N-PSDB; AAS85760.

XX New isolated polynucleotide and encoded polypeptides, useful in

PT diagnostics, forensics, gene mapping, identification of mutations

PT responsible for genetic disorders or other traits and to assess

PT biodiversity

XX Claim 20: SEQ ID No 51932; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and

CC polypeptide (II) sequences. (I) is useful as hybridisation probes,

CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome

CC and gene mapping, and in recombinant production of (II). The

CC polynucleotides are also used in diagnostics as expressed sequence tags

CC for identifying expressed genes. (I) is useful in gene therapy techniques

CC to restore normal activity of (II) or to treat disease states involving

CC (II). (II) is useful for generating antibodies against it, detecting or

CC quantitating a polypeptide in tissue, as molecular weight markers and as

CC a food supplement. (II) and its binding partners are useful in medical

CC imaging of sites expressing (II). (I) and (II) are useful for treating

CC disorders involving aberrant protein expression or biological activity.

CC The polypeptide and polynucleotide sequences have applications in

CC diagnostics, forensics, gene mapping, identification of mutations in

CC responsible for genetic disorders or other traits to assess biodiversity

CC and to produce other types of data and products dependent on DNA and

CC amino acid sequences. ABG00010-ABG30377 represent novel human

CC diagnostic amino acid sequences of the invention.

CC Note: The sequence data for this patent did not appear in the printed

CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published_pct_sequences.

CC Sequence 1510 AA;

Query Match 19.2%; Score 129; DB 22; Length 1510;
 Best Local Similarity 29.9%; Pred. No. 4.9e-05;
 Matches 43; Conservative 17; Mismatches 30; Indels 54; Gaps 8;

QY 6 LMAS-----AGALALLAAPFAQVTPVTELLANPPAGEWISYGOQENYRHSPLRQ 58
 Db 310 ltwgafndpgelngtlisadtpaea-ispvadq-----dwpaygrngqgrfslpkq 360
 QY 59 ITTENVGOLQLVMA-----RGMQPGKV--QVTPLIHGVMTLANPGDVIALDAKTGD 109
 Db 361 lnadnhnlkeavfrtgdvkqpnpgelntevtpi-----kvqd 400
 QY 110 ---LIMEHRR-----QLPNIAATL 124
 Db 401 tpllctahgrlfaaldaemdgiaatl 424

RESULT 13

AAW09614
 ID AAW09614 standard; Protein: 824 AA.

AC AAW09614;

DT 20-MAY-1997 (first entry)

DE Pathogenic Staphylococcus aureus ORF1 gene product.

DE Pathogenicity; vaccine; immunoassay; antibody.

OS Staphylococcus aureus strain NCTC 8325-4 (ISP8).

PN WO9641878-A1.

PD 27-DEC-1996.

PF 13-JUN-1996; 96WO-AU00353.

PR 13-JUN-1995; 95AU-0003507.

PR (AUSU) UNIV AUSTRALIAN NAT.

PI Adham El-Adhami W, Matthea KI, Stewart PR;

DR WPI; 1997-065465/06.

DR N-PSDB; AAT47757.

XX Nucleic acids specific to pathogenic Staphylococcus aureus strains

PT for diagnosis and immunogenic protein prodn.

XX Claim 30; Page 62-65; 88pp; English.

XX An 824-amino acid protein (AAW09614) and 192-amino acid protein

CC (AAW09615) were identified as products of 2 open reading frames in

CC a nucleic acid (AAT47757) isolated from pathogenic Staphylococcus

CC aureus strain ISP8. The 2 gene products are highly hydrophilic

CC and have high surface probability and antigenic index measurements,

CC suggesting that they are surface expressed. The polypeptides,

CC their fragments, and antibodies raised against them can be used in

CC vaccines and as immunoassay reagents for detecting pathogenic S.

CC aureus strains.

XX Sequence 824 AA;

XX Query Match 11.3%; Score 76; DB 18; Length 824;

XX Best Local Similarity 24.8%; Pred. No. 16;

XX Matches 33; Conservative 18; Mismatches 34; Indels 48; Gaps 7;

QY 19 APAFAQVT-----PVTDELLANPPAGEWISYGOQENYRHSPLRQITTENVGOLQLV 70
 Db 683 apgtkvrtgqgkeltitptlknplgtlsgskskelctkdpnlct----- 732

OY 71 WARGMOPGKQVTPLIHD-----GVATLANP--GDVIO-AIDAKT----- 107
 Db 733 -----eygpeltpnrhdefdpklptgekeevpgkpinpetgdvrvpdsvtckypgv 787
 OY 108 -GDLINEHRRLP 119
 Db 788 kgdsive-keelp 799

RESULT 14

ABE71736
 ID ABE71736 standard; Protein; 1229 AA.

AC ABE71736;

DE 26-MAR-2002 (first entry)

DE Drosophila melanogaster polypeptide SEQ ID NO 42000.

KW Drosophila; developmental biology; cell signalling; insecticide; pharmaceutical.

OS Drosophila melanogaster.

PN WO200171042-A2.

PD 27-SEP-2001.

PF 23-MAR-2001; 2001WO-US09231.

PR 23-MAR-2000; 2000US-191637P.

PR 11-JUL-2000; 2000US-0614150.

PA (PEKE) PE CORP NY.

PI Venter JC, Adams M, Li PWD, Myers EW;

DR WPI; 2001-656860/75.

DR N-PSDB; ABL15839.

PT New isolated nucleic acid detection reagent for detecting 1000 or more

PT genes from Drosophila and for elucidating cell signalling and cell-cell

PT interactions -

PS Disclosure; SEQ ID NO 42000; 21pp + Sequence Listing; English.

PS The invention relates to an isolated nucleic acid detection reagent

CC capable of detecting 1000 or more genes from Drosophila. The invention is

CC useful in developmental biology and in elucidating cell signalling and

CC cell-cell interactions in higher eukaryotes for the development of

CC insecticides, therapeutics and pharmaceutical drugs. The invention

CC discloses genomic DNA sequences (ABL16176-AB130511), expressed DNA

CC sequences (ABL16176-AB16175) and the encoded proteins.

CC (AB16176-AB16175) and the encoded proteins.

CC The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 1229 AA;

RESULT 15
 ID AAW35001 standard; Protein; 343 AA.
 AC AAW35001;
 DE 21-MAY-1998 (first entry)
 DE Bacillus thermoovorans endoglucanase.
 KW Endoglucanase; cellulase; carboxymethylcellulose; cellulose;
 KW biomass; beta-1,4-glycosidic bond; hydrolysis; saccharification;
 KW thermostable enzyme; thermophilic; glycosidase.
 OS Bacillus thermoovorans (Clone 68GCl).
 PN WO9744361-A1.
 PD 27-NOV-1997.
 PF 22-MAY-1997; 97MO-US08793.
 PR 22-MAY-1996; 96US-0651572.
 PA (RECO-) RECOMBINANT BIOCATALYSIS INC.
 PI Lam DE, Mathur EJ;
 DR WPI; 1998-018435/02.
 DR N-PSDB; AAT94209.
 PT Endoglucanase(s), preferably form archaeal bacterium, AEP11a -
 PT beta-1,4-glycosidic bonds in cellulose
 PS Claim 1; Fig 1Q; 164pp; English.
 CC This protein comprises an endoglucanase of Bacillus thermoovorans
 CC (Clone 68GCl) that is capable of degrading carboxymethylcellulose
 CC and of hydrolysing the beta-1,4-glycosidic bonds in cellulose. It
 CC has homology to an endoglucanase of archaeobacterium AEP11a (see
 CC AAW34985). It can be produced from native cells or from recombinant
 CC host cells, especially prokaryotic host cells transformed with a
 CC plasmid or virus-derived vector including the endoglucanase DNA
 CC (see AAT94209). 24 Endoglucanases (see AAW34986-W35008) are claimed.
 CC They can be used to degrade cellulose for the conversion of plant
 CC biomass into fuels and chemicals, for use in detergents, textiles,
 CC animal feed, waste treatment, and in the fruit juice and brewing
 CC industries for the clarification and extraction of juices.
 XX Sequence 343 AA;

Query Match 10.8%; Score 72.5; DB 19; Length 343;
 Best Local Similarity 35.3%; Pred. No. 12;
 Matches 18; Conservative 6; Mismatches 26; Indels 1; Gaps 1;

OY 28 VTDELANPPAGEWISYGOENRHSPLTQITTEWGOQLVWARGMOPG 78
 Db 106 vlenasnpiegewekygvitkwestaldattrehgkyrywag-kdpg 155

Search completed: May 24, 2002, 10:05:01
 Job time: 240 sec

OY 8 WMSAGLALLAPAFQOV--TPVTDELANP--PAGEWISYGOEN--YRHSPLTQITTE 62
 Db 958 yvskpavsaasapalakavptvshstgspclgygvgygssnggvl11gpalckltsa 1017
 OY 63 NV-----GOLQVWARGMOPG 78
 Db 1018 payalggkiasstayaigag 1037

Fri May 24 11:27:15 2002

us-08-934-506a-5_copy_1_128.rag

GenCore version 4.5
Copyright (c) 1993 - 2000 Compen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2002, 10:02:56 ; Search time 44.35 Seconds
(without alignments)
70.496 Million cell updates/sec

Title: US-08-934-506a-5_COPY_1_128

Perfect score: 673
Sequence: 1 MKPSTLWASAGALALIAAP.....DLWEHRRLPNTLNSFG 128

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: Issued_Patents_AA:*
2: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	189.5	28.2	754	4	US-09-296-284-4
2	183	27.2	738	1	US-07-985-458-3
3	180.5	26.8	720	4	US-09-296-284-25
4	145.5	21.6	739	3	US-09-136-251-2
5	72.5	10.8	462	4	US-08-630-915A-38
6	70.5	10.5	540	1	US-08-368-071-11
7	70.5	10.5	540	1	US-08-458-181-11
8	70.5	10.5	540	5	PCT-US93-02172-11
9	68	10.1	159	3	US-08-796-792-2
10	67.5	10.0	974	2	US-08-868-786-6
11	65.5	9.7	876	2	US-08-510-215A-2
12	65.5	9.7	1229	3	US-09-310-293-2
13	65.5	9.7	1229	4	US-09-379-376-2
14	65	9.7	851	1	US-08-369-796-2
15	65	9.7	851	2	US-08-852-091-2
16	65	9.7	851	2	US-08-820-754-2
17	65	9.7	851	3	US-08-956-652-2
18	65	9.7	851	3	US-08-956-869-2
19	65	9.7	851	3	US-09-012-710-2
20	65	9.7	851	3	US-08-948-547-2
21	65	9.7	851	4	US-08-087-465-4
22	65	9.7	851	4	US-09-364-970-2
23	65	9.7	851	4	US-09-556-273-2
24	65	9.7	851	4	US-08-956-653A-2
25	65	9.7	851	5	PCT-US95-17025-2
26	65	9.7	852	1	US-08-276-099A-13
27	65	9.7	852	1	US-08-761-890-13

28	64.5	9.6	286	4	US-08-964-127-4	Sequence 4, Appl
29	64.5	9.6	286	4	US-09-496-692-4	Sequence 2, Appl
30	64.5	9.6	520	4	US-08-964-127-2	Sequence 4, Appl
31	64.5	9.6	520	4	US-09-496-692-2	Sequence 2, Appl
32	64.5	9.6	523	2	US-08-923-772-2	Sequence 2, Appl
33	64.5	9.6	523	4	US-09-385-287-2	Sequence 2, Appl
34	64	9.5	966	2	US-08-868-786-2	Sequence 2, Appl
35	63.5	9.4	210	4	US-09-247-155-117	Sequence 11, App
36	63.5	9.4	820	4	US-09-313-677-21	Sequence 21, Appl
37	63.5	9.4	926	4	US-09-313-677-2	Sequence 2, Appl
38	63.5	9.4	933	4	US-09-313-677-19	Sequence 17, Appl
39	63.5	9.4	967	4	US-09-313-677-17	Sequence 3, Appl
40	63	9.4	733	4	US-09-073-587-3	Sequence 1, Appl
41	62.5	9.3	240	4	US-08-913-014A-1	Sequence 4, Appl
42	62.5	9.3	240	4	US-09-072-993C-4	Sequence 12, Appl
43	62	9.2	409	4	US-09-385-028-12	Sequence 3, Appl
44	62	9.2	444	1	US-07-881-075-3	Sequence 3, Appl
45	62	9.2	444	1	US-08-120-827-3	Sequence 3, Appl

ALIGNMENTS

RESULT 1
US-09-296-284-4
; Sequence 4, Application US/09296284A
; Patent No. 6204040
; GENERAL INFORMATION:
; APPLICANT: Choi, Eun-Sung
; APPLICANT: Rhee, Sang-Ki
; APPLICANT: Lee, Eun-Hae
; TITLE OF INVENTION: Gluconobacter Suboxydans Sorbitol Dehydrogenase, Genes
; FILE REFERENCE: 1533.087000
; CURRENT APPLICATION NUMBER: US/09/296,284A
; NUMBER OF SEQ ID NOS: 87
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 754
; TYPE: PRT
; ORGANISM: Gluconobacter suboxydans
US-09-296-284-4

Query Match 28.2%; Score 189.5; DB 4; Length 754;
Best Local Similarity 39.0%; Pred. No. 1e-13;
Matches 46; Conservative 17; Mismatches 46; Indels 9; Gaps 3;
OY 14 LALLAAPAFQVTPV-----TDELLANPAGEWISYCONCENTRHSPLTQITTEWVGO 66
DB 17 LGCMAALAFATSPVALAEDTGTATNADQHPDWMYSYRTYSEQYSPLDITKDNASN 76
OY 67 LQLVNARGMGPQVVOV-TPLIHGVVYLANPGDVIAIDAKGDLWEHRLOLP-NIA 122
DB 77 LKLAHMDLDTFNRGQSGTFLVDGVMTATTNMSKAKALDAATKILMSIDPKVPGNIA 134
RESULT 2
US-07-985-458-3
; Sequence 3, Application US/07985458
; Patent No. 5344777
; GENERAL INFORMATION:
; APPLICANT: Tamaki, Toshimi;
; APPLICANT: Takemura, Hiroshi;
; APPLICANT: Tayama, Kenji;
; APPLICANT: Fukaya, Masahiro;
; APPLICANT: Okumura, Hajime and
; APPLICANT: Kawamura, Yoshiya
; TITLE OF INVENTION: Structural Gene of Membrane-Bound
; TITLE OF INVENTION: Alcohol Dehydrogenase Complex, Plasmid
; TITLE OF INVENTION: Containing The Same And Transformed Acetic Acid
; Bacteria

```

? NUMBER OF SEQUENCES: 4
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Frislauff, Holtz, Goodman & Woodward, P.C.
? STREET: 600 Third Avenue
? CITY: New York
? STATE: New York
? COUNTRY: USA
? ZIP: 10016-2088
? COMPUTER READABLE FORM:
? MEDIUM TYPE: 3.5 inch, 0.72 mb
? COMPUTER: IBM PC compatible (NEC PC-9801 ES)
? OPERATING SYSTEM: MS DOS
? SOFTWARE: ASCII Form
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/07/985,458
? FILING DATE: 19921203
? CLASSIFICATION: 435
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 07/658,221
? FILING DATE: 20-FEB-1991
? APPLICATION NUMBER: 73440/1990
? FILING DATE: 26-MAR-1990
? ATTORNEY/AGENT INFORMATION:
? NAME: Goodman, Herbert
? REGISTRATION NUMBER: 17081
? REFERENCE/DOCKET NUMBER: 910134/HG
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (212)972-1400
? TELEFAX: (212)370-1622
? TELEX: 236268
? INFORMATION FOR SEQ ID NO. 3:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 738 amino acids
? TYPE: amino acid
? TOPOLOGY: unknown
? MOLECULE TYPE: protein
? FEATURE:
? NAME/KEY: MATURE PEPTIDE
? LOCATION: 36 to 738
? IDENTIFICATION METHOD: N-terminal sequences of the
? IDENTIFICATION METHOD: purified protein having a molecular weight of about
? IDENTIFICATION METHOD: 72,000
? ORIGINAL SOURCE:
? ORGANISM: Acetobacter altoacetigenes
? STRAIN: MH-24
? PUBLICATION INFORMATION:
? AUTHORS: Tamaki, Toshiaki;
? AUTHORS: Fukaya, Masahiro;
? AUTHORS: Takemura, Hiroshi;
? AUTHORS: Tayama, Kenji;
? AUTHORS: Okumura, Hajime;
? AUTHORS: Kawamura, Yoshiya;
? AUTHORS: Nishiyama, Makoto;
? AUTHORS: Horinouchi, Sueharu and
? TITLE: Cloning and Sequencing of the Gene Cluster
? TITLE: Encoding Two Subunits of Membrane-Bound
? TITLE: Alcohol Dehydrogenase from Acetobacter
? TITLE: Polyoxygenes
? JOURNAL: Biochimica et Biophysica Acta.
? VOLUME: 1088
? PAGES: 292-300
? DATE: 1991
? US-07-985-458-3

```

Query Match 27.2%; Score 183; DB 1; Length 738;
 Best Local Similarity 33.3%; Pred. No. 5,7e-13;
 Matches 41; Conservative 23; Mismatches 49; Indels 10; Gaps 3;

OY 10 SAGLALIAAPAFQVPTVDE-----LLANPPAGWISYQNGENRHSPLPQIT 61
 DB 17 TAGTICAAALISGVIATMASADGAGATGATGAIHADHPGNNMTYGRVTSIDRYSPLDQINR 76

```

OY 62 ENVGLOLWARGNQPKGVQV-TPLIHDGVWYLANPGDVIAIDAKTGDLIMEHRRLP- 119
DB 77 SNVGNLKIATWLTLDITNRGQEGTPLVIDGVYATTTNMSKMAVADATGKILMSTDPVRVG 136
OY 120 NIA 122
DB 137 NIA 139

```

```

RESULT 3
US-09-296-284-25
? Sequence 25, Application US/09296284A
? Patent No. 6204040
? GENERAL INFORMATION:
? APPLICANT: Choi, Eun-Sung
? APPLICANT: Khee, Sang-Ki
? APPLICANT: Lee, Eun-Hae
? TITLE OF INVENTION: Glucanobacter Suboxydans Sorbitol Dehydrogenase, Genes
? TITLE OF INVENTION: and Methods of Use Thereof
? FILE REFERENCE: 1533.0870000
? CURRENT APPLICATION NUMBER: US/09/296,284A
? CURRENT FILING DATE: 1999-04-22
? NUMBER OF SEQ ID NOS: 87
? SOFTWARE: PatentIn Ver. 2.0
? SEQ ID NO 25
? LENGTH: 720
? TYPE: PRT
? ORGANISM: Glucanobacter suboxydans
US-09-296-284-25

```

```

Query Match 26.8%; Score 180.5; DB 4; Length 720;
Best Local Similarity 40.4%; Pred. No. 1.1e-12;
Matches 40; Conservative 18; Mismatches 36; Indels 5; Gaps 3;
OY 26 TPVTDELLANPPAGWISYQNGENRHSPLQITTEVNGQLQLVARGNQPKGVQV-TP 84
DB 5 TATN--ADQHPGDMWSYGRVTSSEQRYSPLDQITRDNASNKLAWHYLDITNRGQEGTP 61
OY 85 LIHDGVWYLANPGDVIAIDAKTGDLIMEHRRLP-NIA 122
DB 62 LIYDGVMTATTNMSKMAALDAATGKILMSTDPVRPNIA 100

```

```

RESULT 4
US-09-136-251-2
? Sequence 2, Application US/09136251A
? Patent No. 6127156
? GENERAL INFORMATION:
? APPLICANT: HOSHINO, Tatsuo
? APPLICANT: MIYAZAKI, Taro
? APPLICANT: OIYMA, Setsuko
? APPLICANT: SHINOH, Masako
? APPLICANT: TOMIYAMA, No. 61271561bun1
? TITLE OF INVENTION: D-SORBITOL DEHYDROGENASE GENE
? FILE REFERENCE: D-Sorbitol Dehydrogenase Gene
? CURRENT APPLICATION NUMBER: US/09/136,251A
? CURRENT FILING DATE: 1998-08-19
? EARLIER APPLICATION NUMBER: EP 97114432.4
? EARLIER FILING DATE: 1997-08-21
? NUMBER OF SEQ ID NOS: 9
? SOFTWARE: PatentIn Ver. 2.1
? SEQ ID NO 2
? LENGTH: 739
? TYPE: PRT
? ORGANISM: Glucanobacter suboxydans
? FEATURE:
? NAME/KEY: SIGNAL
? LOCATION: (1)..(24)
US-09-136-251-2

```


RESULT 8
 PCT-US93-02172-11
 Sequence 11: Application PC/TUS9302172
 GENERAL INFORMATION:
 APPLICANT: La Jolla Cancer Research Foundation
 TITLE OF INVENTION: RECOMBINANT CALF INTESTINAL ALKALINE
 PHOSPHATASE
 NUMBER OF SEQUENCES: 13
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: La Jolla Cancer Research Foundation
 STREET: 10901 North Torrey Pines Road
 CITY: La Jolla
 STATE: California
 COUNTRY: USA
 ZIP: 92037
 COMPUTER READABLE FORM:
 MEDIUM TYPE: floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version 1.25 (EPO)
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: PCT/US93/02172
 FILING DATE: 19930310
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/07/849,219
 FILING DATE: 10-MAR-1992
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (619) 455-6480
 TELEFAX: (619) 455-0181

ZAP: 10111
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: Windows95
 SOFTWARE: Fastset for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/796,792
 FILING DATE: 06-FEB-1997
 CLASSIFICATION: 536
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 60/011,364
 FILING DATE: 09-FEB-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: Hone, William J.
 REGISTRATION NUMBER: 26,739
 REFERENCE/DOCKET NUMBER: 07763/03301
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 212-765-5070
 TELEFAX: 212-258-2291
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 159 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 FRAGMENT TYPE: Internal
 S-08-796-792-2


```

; Patent No. 6117662
; GENERAL INFORMATION:
; APPLICANT: Warren, Richard L.
; APPLICANT: Van Horn, Stephanie
; APPLICANT: Traini, Christopher M.
; APPLICANT: Mathie, Thomas B.
; TITLE OF INVENTION: Respiratory nitrate reductase alpha
; FILE REFERENCE: subunit
; CURRENT APPLICATION NUMBER: US/09/310,293
; CURRENT FILING DATE: 1999-05-12
; EARLIER APPLICATION NUMBER: 60/086,579
; EARLIER FILING DATE: 1998-05-22
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1229
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-310-293-2

```

```

Query Match
Best Local Similarity 9.7%; Score 65.5; DB 3; Length 1229;
Matches 33; Conservative 22; Mismatches 62; Indels 25; Gaps 8;

```

```

OY 2 KPTSLM--ASAGALALAAPAFQ-VTPVDELLANPPAGWIS---YQNGENYRH- 53
DB 739 KPEITWRETTGKLDIVVSLDFRTATPLSDIVL--PAATWYKHDLSSTDMHPYVHP 796
OY 54 -----SPLTQITTE-----NVGQLQVWARGMOPG---KVQVPLIHGVTYLANPGDVI 100
DB 797 FNPALDPLMESRSDMDIYKTLAKAFSEMAKDYLPGTFRKDVYVTPPLSHDTKQELSTPYGVV 856
OY 101 QAIADKTDGLMEHRRQLPNTA 122
DB 857 K--DMSKGEIAPVPGRTMPNFA 876

```

```

RESULT 13
US-09-579-376-2
; Sequence 2, Application US/09579376
; Patent No. 6316237
; GENERAL INFORMATION:
; APPLICANT: Warren, Richard L.
; APPLICANT: Van Horn, Stephanie
; APPLICANT: Traini, Christopher M.
; APPLICANT: Mathie, Thomas B.
; TITLE OF INVENTION: Respiratory nitrate reductase alpha
; FILE REFERENCE: subunit
; CURRENT APPLICATION NUMBER: US/09/579,376
; CURRENT FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 09/310,293
; PRIOR FILING DATE: 1999-05-12
; PRIOR APPLICATION NUMBER: 60/086,579
; PRIOR FILING DATE: 1998-05-22
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1229
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-579-376-2

```

```

Query Match
Best Local Similarity 9.7%; Score 65.5; DB 4; Length 1229;
Matches 33; Conservative 22; Mismatches 62; Indels 25; Gaps 8;
OY 2 KPTSLM--ASAGALALAAPAFQ-VTPVDELLANPPAGWIS---YQNGENYRH- 53
DB 739 KPEITWRETTGKLDIVVSLDFRTATPLSDIVL--PAATWYKHDLSSTDMHPYVHP 796

```

```

OY 54 -----SPLTQITTE-----NVGQLQVWARGMOPG---KVQVPLIHGVTYLANPGDVI 100
DB 797 FNPALDPLMESRSDMDIYKTLAKAFSEMAKDYLPGTFRKDVYVTPPLSHDTKQELSTPYGVV 856
OY 101 QAIADKTDGLMEHRRQLPNTA 122
DB 857 K--DMSKGEIAPVPGRTMPNFA 876

```

```

RESULT 14
US-08-369-796-2
; Sequence 2, Application US/08369796
; Patent No. 5716622
; GENERAL INFORMATION:
; APPLICANT: James E. Darnell, Jr.
; APPLICANT: Zilong Wen
; APPLICANT: Curt M. Horvath
; APPLICANT: Zhong Zhong
; TITLE OF INVENTION: FUNCTIONALLY ACTIVE REGIONS OF SIGNAL
; NUMBER OF INVENTIONS: TRANSDUCER AND ACTIVATOR OF TRANSCRIPTION (STAT) PROTEINS
; CORRESPONDENCE ADDRESS:
; ADDRESS: Klauber & Jackson
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/369,796
; FILING DATE: 06-JAN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-116
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201 487-5800
; TELEFAX: 201 343-1684
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 851 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-369-796-2

```

```

Query Match
Best Local Similarity 9.7%; Score 65; DB 1; Length 851;
Matches 30; Conservative 16; Mismatches 43; Indels 42; Gaps 6;
OY 1 MKPTSLMASAGALALALAAPAFQVTPVDELLANPPAGW-----ISYQ 46
DB 459 MNGSLAMASVLMFNLSPNLOQ-----QFESNPKAPWSLGLPALSMQFSSYVGRGL 512
OY 47 NOENRHSPLTQITTEYNGQ-----LQLYVA-----RGMOPGKV-----QVPLIHG 89
DB 513 NSDQ-----LSMLRNKLFEGQNCRTEDPLLSWADFTKRESPPGKLPFTWIDKILELVHDH 567
OY 90 VMYLANPGDVI 100
DB 568 LKDLMDGRIM 578

```

```

RESULT 15

```


GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2002, 10:06:03 ; Search time 56.04 Seconds
(without alignments)
219.476 Million cell updates/sec

Title: US-08-934-506A-5_COPY_1_128
Sequence: 1 MKPTSLIMASGALALLAAP.....DLWEHRRLPNTATINSTE 128

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues
Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

PIR_71:*
1: PIR1:*
2: PIR2:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	190	28.2	742	2 A49340	alcohol dehydrogen
2	183	27.2	738	2 S14270	alcohol dehydrogen
3	178	26.4	742	2 JS0326	alcohol dehydrogen
4	165.5	24.6	708	2 S52317	quinohemoprotein e
5	155	23.0	796	1 JV0107	glucose dehydrogen
6	155	23.0	796	2 H90644	glucose dehydrogen
7	155	23.0	796	2 H85495	glucose dehydrogen
8	154	22.9	796	2 AG0523	glucose dehydrogen
9	148	22.0	778	2 G98221	glucose dehydrogen
10	148	22.0	778	2 A13064	glucose dehydrogen
11	137.5	20.4	803	2 B83360	glucose dehydrogen
12	137	20.4	623	2 B83399	glucose dehydrogen
13	134.5	20.0	601	2 E95863	glucose dehydrogen
14	134	19.9	808	1 QPKEX	glucose dehydrogen
15	121.5	18.1	801	1 S00943	glucose dehydrogen
16	106.5	15.8	809	2 A12968	glucose dehydrogen
17	106.5	15.8	809	2 A12968	glucose dehydrogen
18	101.5	15.1	626	2 JQ0706	glucose dehydrogen
19	94.5	14.0	809	2 A55547	glucose dehydrogen
20	93.5	13.9	639	2 JC4881	glucose dehydrogen
21	83.5	12.4	2802	2 F97686	glucose dehydrogen
22	83.5	12.4	2831	2 A12911	glucose dehydrogen
23	82	12.2	668	2 C75264	glucose dehydrogen
24	81.5	12.1	393	2 AD0350	glucose dehydrogen
25	81	12.0	613	2 F69424	glucose dehydrogen
26	80.5	12.0	647	2 C83012	glucose dehydrogen
27	78.5	11.7	509	2 S05638	glucose dehydrogen
28	78.5	11.7	509	2 C70624	glucose dehydrogen
29	77	11.4	733	2 S78376	glucose dehydrogen

30	76.5	11.4	224	2 AD2115	phosphoribosylform
31	76.5	11.4	400	2 AD0604	D-alanyl-D-alanine
32	76	11.3	306	2 D75562	N-acetyl-gamma-glu
33	76	11.3	386	2 A82284	conserved hypotet
34	76	11.3	573	2 S68591	methanol dehydrog
35	76	11.3	580	2 A80394	gamma-glutamyltran
36	76	11.3	772	2 H86016	hypothetical prote
37	75.5	11.2	407	2 C82804	conserved hypotet
38	75.5	11.2	684	2 T36771	hypothetical prote
39	75.5	11.2	3972	2 S75251	conserved hypotet
40	74	11.0	380	2 C83171	ser/thr protein k1
41	74	11.0	969	2 A75634	McR-related prote
42	73.5	10.9	584	2 A82506	probable sugar upt
43	73	10.8	355	2 E95990	probable lipoprote
44	73	10.8	392	2 AG0821	secreted protein,
45	73	10.8	407	2 G87515	

ALIGNMENTS

RESULT 1

A49340
alcohol dehydrogenase (EC 1.1.1.-) precursor - Acetobacter pasteurianus (strain NC113
C:Species: Acetobacter pasteurianus
C:Date: 07-Apr-1994 #sequence_revision 07-Apr-1994 #text_change 20-Jun-2000
C:Accession: A49340
R:Takemura, H.; Kondo, K.; Horinouchi, S.; Beppu, T.
J. Bacteriol. 175, 6857-6866, 1993
A:Title: Induction by ethanol of alcohol dehydrogenase activity in Acetobacter paste
A:Reference number: A49340; MUID:94042848
A:Accession: A49340
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-742 <TAK>
A:Cross-references: GB:D13893; NID:95117067; PIDN:BAA0252.1; PID:9452586
C:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)
C:Keywords: alcohol metabolism; NAD; oxidoreductase

Query Match

Best Local Similarity 38.8%; Pred. No. 4.9e-11;
Matches 45; Conservative 18; Mismatches 49; Indels 4; Gaps 3;

QY 11 AGALLAAPAAQVPTVDLL--ANPPAGEMISYCGNENYRHSPLQITENVGOLQ 68
DB 23 AALPFAAIPARADGGGNTGAEIIRHDDHPENWLSYGRYSQRYSPLDQINRSNVGLK 82
QY 69 LWMARGMPGKYV-VPTLHDGMYLANPGDYIQADTGDILWEHRRLP-NTA 122
DB 83 LAMVYTLIDNRGGENTPLVVDGIMYATTKMSKEMALDATGKLWQYDKVGNIA 138

RESULT 2

S14270
alcohol dehydrogenase, membrane-bound (EC 1.1.1.-) 72K chain precursor - Acetobacter
C:Species: Acetobacter pasteurianus
A:Variety: strain NB11028
C:Date: 21-Nov-1993 #sequence_revision 26-Jul-1996 #text_change 20-Jun-2000
C:Accession: S14270
R:Tamaki, T.; Fukaya, M.; Takemura, H.; Tayama, K.; Okumura, H.; Kawamura, Y.; Nishiy
Biochim. Biophys. Acta 1088, 292-300, 1991
A:Title: Cloning and sequencing of the gene cluster encoding two subunits of membrane
A:Reference number: S14270; MUID:91159482
A:Accession: S14270
A:Molecule type: DNA
A:Residues: 1-738 <TAM>
A:Cross-references: GB:D00635; NID:9216185; PIDN:BAA00528.1; PID:9216186
A:Experimental source: strain NB11028
C:Complex: heterodimer of 72K and 44K chains
C:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)
C:Keywords: alcohol metabolism; membrane protein; oxidoreductase
F:1-35/Domain: signal sequence #status predicted <SIG>

F:36-738/Product: alcohol dehydrogenase 72k chain #status predicted <MAY>

Query Match 27.2%; Score 183; DB 2; Length 738;
Best Local Similarity 33.3%; Pred. No. 2.5e-10;
Matches 41; Conservative 23; Mismatches 49; Indels 10; Gaps 3;

QY 10 SAGALALAAPAFACVTPVTDE-----LLANPPAGWISYSGONGENRHSPLTQITTT 61
DB 17 TAGTICALISGVATFASADSDGATGEGAIHADHPGMMWRTGRTYSQRTSPLDQJNR 76
QY 62 ENVGQQLVWARGMOPGKQV-TPLIHGVMYLANPGDVIQDAIKTGLIWEHRRLP- 119
DB 77 SHVGNKILAMYLDDLTNRQEGTPLVIDGVMTATMSMKAVDAATKLLMSYDPVPG 136
QY 120 NIA 122
DB 137 NIA 139

RESULT 3

alcohol dehydrogenase (EC 1.1.-.-) 72k chain precursor - Acetobacter aceti
C:Species: Acetobacter aceti
C:Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 20-Jun-2000
C:Accession: J50326
R:Inoue, T.; Sunagawa, M.; Mori, A.; Imai, C.; Fukuda, M.; Takagi, M.; Yano, K.
J. Bacteriol. 171, 3115-3122, 1989
A:Title: Cloning and sequencing of the gene encoding the 72-kilodalton dehydrogenase subunit
A:Reference number: J50326; MUID:89255070
A:Accession: J50326
A:Molecule type: DNA
A:Residues: 1-742 <IMO>
A:Cross-References: GB:D90004; GB:M26951; NID:g216193; PIDN:BAAI4058.1; PID:g216194
A:Experimental source: strain K6033
A:Note: amino terminal of mature protein is confirmed
C:Genetics:
A:Gene: adhI
C:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)
C:Keywords: alcohol metabolism; NAD; oxidoreductase
F:1-35/Domain: signal sequence #status predicted <SIG>
F:36-742/Product: alcohol dehydrogenase 72k dehydrogenase chain #status predicted <ADC>

Query Match 26.4%; Score 178; DB 2; Length 742;
Best Local Similarity 37.9%; Pred. No. 7.9e-10;
Matches 44; Conservative 19; Mismatches 49; Indels 4; Gaps 3;

QY 11 AGALALAAPAFACVTPVTDEL--ANPPAGWISYSGONGENRHSPLTQITTEVGOLO 68
DB 23 AALPFAAVPARADGCGTGEAIHDDHDENNLSYGRITSEGRYSPLDQJNSVGDLE 82
QY 69 LVNARGQPGKQV-TPLIHGVMYLANPGDVIQDAIKTGLIWEHRRLP-NIA 122
DB 83 LGGYTLDTNRQEGTPLVIDGVMTATMSMKALDAATGKLLMQYDPRVGNIA 138

RESULT 4

quinohemoprotein ethanol dehydrogenase (EC 1.1.99.-) type 1 precursor - Comamonas testos
C:Species: Comamonas testosteroni
C:Date: 08-May-1995 #sequence_revision 21-Jul-1995 #text_change 02-Jun-2000
C:Accession: S62366; S62373; S65908; S52317
R:Stoorvogel, J.; Kraayveld, D.E.; van Sluis, C.A.; Jongejan, J.A.; de Vries, S.; Duine, Eur. J. Biochem. 235, 690-698, 1996
A:Title: Characterization of the gene encoding quinohemoprotein ethanol dehydrogenase C
A:Reference number: S62366; MUID:96184549
A:Accession: S62366
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-708 <STOI>
A:Cross-References: EMBL:X81880; NID:g663195; PIDN:CAA57464.1; PID:g663196
A:Accession: S62373

A:Status: preliminary
A:Molecule type: protein
A:Residues: 32-54 <STOI>
R:de Jong, G.A.H.; Geerlof, A.; Stoorvogel, J.; Jongejan, J.A.; de Vries, S.; Duine, Eur. J. Biochem. 230, 899-905, 1995
A:Title: Quinohemoprotein ethanol dehydrogenase from Comamonas testosteroni. Purific
A:Reference number: S65908; MUID:95324380
A:Accession: S65908
A:Molecule type: protein
A:Residues: 32-50, 'X', 52-54; 477-483, 'X', 485-490 <DEJ>
A:Experimental source: ATCC 15667
C:Genetics:
A:Gene: qheh
C:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)
C:Keywords: oxidoreductase; quinoprotein
F:1-31/Domain: signal sequence #status predicted <SIG>
F:32-708/Product: quinohemoprotein ethanol dehydrogenase type 1 #status experimental

Query Match 24.6%; Score 165.5; DB 2; Length 708;
Best Local Similarity 31.8%; Pred. No. 1.4e-08;
Matches 41; Conservative 23; Mismatches 50; Indels 15; Gaps 4;

QY 3 PLSLWASAGALALAAPAFACVTPVT-----DELLANPP-AGWISYSGONGEN 50
DB 12 PGRWVWLLAACLG--SAAFAQGTGPAQAQAAVORVDSGFIRANARTDWPFGVYAE 69
QY 51 YRSPSLTQITTEVGOLOLVWARGMOPGK-VQVTPLIHGVWYLANPGDVIQDAIKTGD 109
DB 70 TRYSRLDQINAAVWKLGLAMSYNESTRGVEATPVVDGIMVVSASVVAIDTRIGN 129
QY 110 LIWEHRRL 118
DB 130 RIWTPDQI 138

RESULT 5

glucose dehydrogenase (pyrroloquinoline-quinone) (EC 1.1.99.17) - Escherichia coli
C:Species: Escherichia coli
C:Date: 30-Jun-1992 #sequence_revision 12-Dec-1997 #text_change 16-Jun-2000
C:Accession: D64735; J50107; A45997; S45201; I41228
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.;
A.: Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID:97426617
A:Accession: D64735
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-796 <BLAT>
A:Cross-References: GB:A6000122; GB:U00096; NID:g1786315; PIDN:AACT3235.1; PID:g17863
A:Experimental source: strain K-12, substrain MG1655
R:Cleton-Jansen, A.M.; Goosen, N.; Fayet, O.; van de Putte, P.
J. Bacteriol. 172, 6308-6315, 1990
A:Title: Cloning, mapping, and sequencing of the gene encoding Escherichia coli quino
A:Reference number: J50107; MUID:91035240
A:Accession: J50107
A:Molecule type: DNA
A:Residues: 1-58, 'L', 60-148, 'H', 150, 'KRRCHT', 157-192, 'K', 194-796 <CLE>
A:Cross-References: GB:X51323; NID:g41553; PIDN:CAA35706.1; PID:g41554
A:Experimental source: strain K12
R:Yamada, M.; Sumi, K.; Matsushita, K.; Adachi, O.; Yamada, Y.
J. Biol. Chem. 268, 12812-12817, 1993
A:Title: Topological analysis of quinoprotein glucose dehydrogenase in Escherichia co
A:Reference number: A45997; MUID:93286127
A:Accession: A45997
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-20 <YAM>
R:Fujita, N.
submitted to the EMBL Data Library, January 1994
A:Reference number: S45181

A:Residues: 1-801 <CLP>
A:Cross-references: EMBL:X07235; NID:g38711; PIDN:CAA30222.1; PID:g38712
A:Experimental source: strain LMD 79.41
C:Function:
A:Description: catalyzes the oxidation of D-glucose to D-gluconic acid by ubiquinone
A:Pathway: respiratory chain
C:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)
C:Keywords: oxidoreductase; pyrroloquinoline quinone; respiratory chain; transmembrane F
F:9-35/Domain: transmembrane #status predicted <TM1>
F:39-57/Domain: transmembrane #status predicted <TM2>
F:61-79/Domain: transmembrane #status predicted <TM3>
F:94-108/Domain: transmembrane #status predicted <TM4>
F:118-137/Domain: transmembrane #status predicted <TM5>
F:91_93/Binding site: ubiquinone (Arg, Asp) #status predicted
F:471/Active site: Asp #status predicted

Query Match 18.1%; Score 121.5; DB 1; Length 801;
Best Local Similarity 30.3%; Pred. No. 0.00043;
Matches 36; Conservative 15; Mismatches 45; Indels 23; Gaps 4;
QY 20 PAFQVTPYTDLLANPPAGEMISYGQNDENRHSPLQTITENVGQLQVWARGMOPGK 79
DB 152 PETAQAVPGVAE-----SDMPAYGRTOAGVRYSPKQINDQNVKDLKVAWT--LRTGD 202
QY 80 V-----QVTPLIHGVMYLANPGDVYQAIDAKTGDLWEHRROLPNIAITLNSF 127
DB 203 LKTDNDSGFTNQVPIKIGNMNFICTAHQOLAIIDPATGKEKW---RFDPKRLKTDKSF 258

Search completed: May 24, 2002, 10:06:04
Job time: 288 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2002, 10:15:48 ; Search time 27.86 Seconds
(without alignments)
177.893 Million cell updates/sec

Title: US-08-934-506a-5_COPY_1_128

Perfect score: 673
Sequence: 1 MKPTSLIMASAGALALAAP.....DLIWEHRQLPNIATLNSRG 128

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SWISSPROT_40:*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	ID	Description
1	206	30.6	757	1 DHET_GLUOX	005542 gluconobact
2	183	27.2	738	1 DHET_ACEPO	P28036 acetobacter
3	182	27.0	739	1 DHET_ACEEU	Q44002 acetobacter
4	178	26.4	742	1 DHET_ACEAC	P18278 acetobacter
5	155	23.0	796	1 DHG_ECOLI	P15877 escherichia
6	137	20.4	623	1 EXAA_PSEAE	Q92417 pseudomonas
7	134	19.9	808	1 DHG_GLUOX	P27175 gluconobact
8	121.5	18.1	801	1 DHGA_ACICA	P05465 acinetobact
9	112.5	16.7	631	1 DHM1_PARDE	P12293 paracoccus
10	111.5	16.6	626	1 DHM1_METOR	P15279 methyllobact
11	101.5	15.1	626	1 DHM1_METEX	P16027 methyllobact
12	94.5	14.0	809	1 QUIA_ACICA	Q59086 acinetobact
13	93	13.8	600	1 XOXF_PARDE	P29968 paracoccus
14	90	13.4	790	1 QUIA_XANCI	Q92478 xanthomonas
15	78.5	11.7	148	1 CU08_LOCOM	P11734 locusta mig
16	77	11.4	733	1 PSAB_OOSI	P19480 odontella s
17	76	11.3	233	1 DHM1_MARMO	Q35734 marmota mon
18	76	11.3	571	1 DHM1_MARME	P38539 methyllobact
19	73	10.8	776	1 PSAB_HETTR	Q92478 xanthomonas
20	70.5	10.5	540	1 PSAB_HETTR	P15693 rattus norv
21	70	10.4	441	1 GUNN_THIEFU	P26222 thermomonas
22	70	10.4	753	1 YBHU_ECOLI	P75764 escherichia
23	69.5	10.3	400	1 DACC_ECOLI	P08506 escherichia
24	69	10.3	640	1 INVI_CAVAN	P93761 capsicum an
25	68.5	10.2	353	1 OM51_HAELIN	P43840 haemophilus
26	68.5	10.2	417	1 HEM2_SELMA	Q04623 selaginella
27	68.5	10.2	524	1 P60_LISIV	P01837 listeria iv
28	68	10.1	159	1 MP63_MCTTU	P97175 mycobacteri
29	68	10.1	347	1 YAB9_MYCPN	P75593 mycoplasma
30	68	10.1	661	1 PSAB_PROHO	P58387 prochloroth
31	68	10.1	1515	1 GUTB_AZOBR	O05755 azospirillum
32	67.5	10.0	146	1 AZUP_ALCFA	P04377 alcaligenes
33	67.5	10.0	261	1 TPIS_MYCLE	P46711 mycobacteri

34	67.5	10.0	734	1 PSAB_CYACA	Q91166 cyanidium c
35	67.5	10.0	974	1 PH52_SOLTY	P53535 solanum tub
36	67	10.0	331	1 DIV_ECOLI	P15286 escherichia
37	67	10.0	485	1 BIND_STREN	P23118 strongyloce
38	67	10.0	500	1 YFJ1_YEAST	P43601 saccharomyc
39	67	10.0	864	1 STR2_PIG	O02799 sus scrofa
40	66.5	9.9	1550	1 GLTB_STN3	P55037 synechocyst
41	66	9.8	346	1 GPDA_XYLEA	O9pcht xyella fas
42	66	9.8	369	1 YDHR_ECOLI	P7570 escherichia
43	66	9.8	527	1 YF00_MYCPN	P75287 mycoplasma
44	66	9.8	755	1 TRZN_AGRVI	P25017 agrobacteri
45	65.5	9.7	207	1 DSBA_YERPE	Q92472 yersinia pe

ALIGNMENTS

RESULT	ID	Query	Match	Score	Description
1	DHET_GLUOX	STANDARD:	757 AA.		
AC	005542				
DT	01-NOV-1997 (Rel. 35, Created)				
DT	01-NOV-1997 (Rel. 35, Last sequence update)				
DT	16-OCT-2001 (Rel. 40, Last annotation update)				
DE	Alcohol dehydrogenase [acceptor] precursor (EC 1.1.99.8) (G3-ADH subunit 1).				
GN	ADHA.				
OS	Gluconobacter oxydans (Gluconobacter suboxydans).				
OC	Bacteria; Proteobacteria; alpha subdivision; Acetobacteraceae;				
OC	Gluconobacter.				
OX	NCBI_Taxid=442;				
RN	[1]				
RP	SEQUENCE FROM N.A., AND SEQUENCE OF 36-50.				
RC	STRAIN-IFO 12528;				
RX	MEDLINE=97208225; PubMed=9055427;				
RA	Kondo K., Horinouchi S.;				
RT	"Characterization of the genes encoding the three-component membrane-bound alcohol dehydrogenase from Gluconobacter suboxydans and their expression in Acetobacter pasteurianus.";				
RT	Appl. Environ. Microbiol. 63:1131-1138(1997).				
RL	Appl. Environ. Microbiol. 63:1131-1138(1997).				
CC	-1- FUNCTION: DEHYDRATION OF PRIMARY ALCOHOLS (EXCEPTION: METHANOL) (BY SIMILARITY).				
CC	-1- CATALYTIC ACTIVITY: A primary alcohol + acceptor = an aldehyde + reduced acceptor.				
CC	-1- COFACTOR: PO4 AND HEME.				
CC	-1- SUBUNIT: HETEROPOLYMER (DEHYDROGENASE AND PROTEIN ADS). THAT FORMS THE ALCOHOL DEHYDROGENASE COMPLEX.				
CC	-1- SUBCELLULAR LOCATION: MEMBRANE-BOUND, FACING THE PERIPLASMIC SPACE (POTENTIAL).				
CC	-1- SIMILARITY: BELONGS TO THE BACTERIAL POQ DEHYDROGENASE FAMILY.				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@sib-sib.ch).				
CC	EMBL: D86375; BAA19753.1; -				
DR	HSSP: Q924U7; IFTG.				
DR	InterPro: IPR001479; Bac_POO.				
DR	InterPro: IPR002372; Bac_POO_repeat.				
DR	InterPro: IPR000345; CytC_heme_bind.				
DR	Pfam: PF01011; Bacterial_POO_6.				
DR	PROSITE: PS00363; BACTERIAL_POO_1; 1.				
DR	PROSITE: PS00364; BACTERIAL_POO_2; 1.				
DR	PROSITE: PS00190; CYTOCHROME_C_1.				
KW	Oxidoreductase; POQ; Heme; Periplasmic; Membrane; Signal.				
FT	SIGNAL	1	34		
FT	CHAIN	35	757		ALCOHOL DEHYDROGENASE [ACCEPTOR].
FT	MOD_RES	35	35		PYRROLIDONE CARBOXYLIC ACID.
FT	ACT_SITE	342	342		BASE (POTENTIAL).

FT BINDING 653 653 HEME (COVALENT) (BY SIMILARITY).
 FT BINDING 656 656 HEME (COVALENT) (BY SIMILARITY).
 FT METAL 657 657 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
 FT METAL 657 657 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
 SO SEQUENCE 757 AA; 82968 MW; 39B9F90EB947581 CRC64;

Query Match
 Best Local Similarity 43.1%; Score 206; DB 1; Length 757;
 Matches 53; Conservative 19; Mismatches 43; Indels 8; Gaps 5;

QY 7 LMSAGALLALIA--PAFAQV---TPVDELLANPPAGWISYSGONQENYRHSPLQITTT 61
 Db 16 LLSGAAALAFSAVPAVPAEDPTGTATSSDNGGHP-GDWLSYGRSYSEORSPDLQINT 74
 QY 62 ENVGQLQVWARGMOPGKQV--TPLIHDSVYLANPBGVIAQDAKTDGLMEHRRLP- 119
 Db 75 ENVGKLKLMHYDLDTNRGQEGTPLVNSVMATITNMSMKRALDAATKGLMSYDPKVP 134
 QY 120 NIA 122
 Db 135 NIA 137

RESULT 2

DHET_ACEPO STANDARD: PRT; 738 AA.

AC P28036; 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Alcohol dehydrogenase [acceptor] precursor (EC 1.1.99.8).
 GN ADHA.
 OS Acetobacter polyoxogenes.
 CC Bacteria; Proteobacteria; alpha subdivision; Acetobacteraceae;
 CC Acetobacter.
 OX NCBI_Taxid=439;
 RP SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.
 RC STRAIN=NB11028;
 RX MEDLINE=91159482; Pubmed=2001402;
 RA Tanaka T., Fukaya M., Takemura H., Tayama K., Okumura H.,
 Kawamura Y., Nishiyama M., Horinouchi S., Beppu T.;
 RT "Cloning and sequencing of the gene cluster encoding two subunits of
 membrane-bound alcohol dehydrogenase from Acetobacter polyoxogenes";
 RL Biochim. Biophys. Acta 1088:292-300(1991).

CC -1- FUNCTION: DEHYDRATION OF PRIMARY ALCOHOLS (EXCEPTION: METHANOL).
 CC -1- CATALYTIC ACTIVITY: A primary alcohol + acceptor = an aldehyde +
 CC reduced acceptor.
 CC -1- COFACTOR: POQ AND HEME.
 CC -1- SUBUNIT: HETEROTETRAMER (DEHYDROGENASE, CYTOCHROME, AND TWO
 CC SMALLER UNKNOWN SUBUNITS) THAT FORMS THE ALCOHOL DEHYDROGENASE
 CC COMPLEX.
 CC -1- SUBCELLULAR LOCATION: MEMBRANE-BOUND, FACING THE PERIPLASMIC
 CC SPACE (POTENTIAL).
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POQ DEHYDROGENASE FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

DR EMBL: D00635; BAA00528.1; -
 DR PIR: S14270; S14270.
 DR HSSP: Q9Z4J7; 1FT6.
 DR InterPro: IPR001479; Bac_POQ.
 DR InterPro: IPR002372; Bac_POQ_repeat.
 DR InterPro: IPR000345; CYC_heme_bind.
 DR Pfam: PF01011; Bacterial_POQ_6.
 DR PROSITE: PS00363; BACTERIAL_POQ_1; 1.
 DR PROSITE: PS00364; BACTERIAL_POQ_2; 1.

DR PROSITE: PS00190; CYTOCHROME_C; 1.
 KW Oxidoreductase; POQ; Heme; Periplasmic; Membrane; Signal.
 FT SIGNAL 1 35
 FT CHAIN 36 738
 FT BINDING 650 650 ALCOHOL DEHYDROGENASE [ACCEPTOR].
 FT BINDING 653 653 HEME (COVALENT) (BY SIMILARITY).
 FT BINDING 653 653 HEME (COVALENT) (BY SIMILARITY).
 FT METAL 654 654 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
 FT METAL 654 654 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
 SO SEQUENCE 738 AA; 80840 MW; 1E2B6ED7BCD92AF6 CRC64;

Query Match
 Best Local Similarity 33.3%; Score 183; DB 1; Length 738;
 Matches 41; Conservative 23; Mismatches 49; Indels 10; Gaps 3;

QY 10 SAGALLALIAFAFAQVTPVDE-----LLANPPAGWISYSGONQENYRHSPLQITTT 61
 Db 17 TAGTICALISGYATMASADGCGATGEATIHADHPGNNMTGGRYSQDRSPDLQINR 76
 QY 62 ENVGQLQVWARGMOPGKQV--TPLIHDSVYLANPBGVIAQDAKTDGLMEHRRLP- 119
 Db 77 SNGNKLKLMHYDLDTNRGQEGTPLVNSVMATITNMSMKRALDAATKGLMSYDPKVP 136
 QY 120 NIA 122
 Db 137 NIA 139

RESULT 3

DHET_ACEEU STANDARD: PRT; 739 AA.

AC Q44002; 007952; 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Alcohol dehydrogenase [acceptor] precursor (EC 1.1.99.8).
 GN ADH.
 OS Acetobacter europaeus.
 CC Bacteria; Proteobacteria; alpha subdivision; Acetobacteraceae;
 CC Gluconacetobacter.
 OX NCBI_Taxid=33995;
 RP SEQUENCE FROM N.A.
 RC STRAIN=DES11 / DSM 6160;
 RA Thurner C.A.K.;
 RT Submitted (NOV-1996) to the EMBL/Genbank/DBJ databases.

CC -1- CATALYTIC ACTIVITY: A primary alcohol + acceptor = an aldehyde +
 CC reduced acceptor.
 CC -1- COFACTOR: POQ AND HEME (BY SIMILARITY).
 CC -1- SUBUNIT: HETEROTETRAMER (DEHYDROGENASE, CYTOCHROME, AND TWO
 CC SMALLER UNKNOWN SUBUNITS) THAT FORMS THE ALCOHOL DEHYDROGENASE
 CC COMPLEX (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: MEMBRANE-BOUND, FACING THE PERIPLASMIC
 CC SPACE (POTENTIAL).
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POQ DEHYDROGENASE FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

DR EMBL: X82894; CAA58066.1; -
 DR PIR: Y09480; CAA70688.1; -
 DR HSSP: Q9Z4J7; 1FT6.
 DR InterPro: IPR001479; Bac_POQ.
 DR InterPro: IPR002372; Bac_POQ_repeat.
 DR InterPro: IPR000345; CYC_heme_bind.
 DR Pfam: PF01011; Bacterial_POQ_6.
 DR PROSITE: PS00363; BACTERIAL_POQ_1; 1.
 DR PROSITE: PS00364; BACTERIAL_POQ_2; 1.
 DR PROSITE: PS00190; CYTOCHROME_C; 1.


```

KW Oxidoreductase; POQ; Heme; Periplasmic; Membrane; Signal.
FT SIGNAL 1 35
FT CHAIN 36 739
FT BINDING 651 651
FT BINDING 654 654
FT METAL 655 655
FT SEQUENCE 739 AA; 80944 MW; E681BB237ACB91F4 CRC64;
SO
Query Match 27.08; Score 182; DB 1; Length 739;
Best Local Similarity 33.38; Pred. No. 2,8e-10;
Matches 41; Conservative 22; Mismatches 50; Indels 10; Gaps 3;
OY 10 SAGMALALLAAFAFYQVPTDE-----LLANPAGEWISYGCQDENVRRSPLQIQT 61
DB 17 TAGTTCALLISGYTWTMSADDDGGATGEALIIHADHGGNMATGGRYSEGRVSPLDQINR 76
OY 62 ENVGGLQLVWARGMQPGKQVY-TPLIHDGWTYLANPGDYIQAIDAKTGLIWEHRQLP- 119
DB 77 SNVGNLKLAWYLIDIDTRNGDETPYLVIGVYATTNMKNKAVDAATGLKMSYDPRVQ 136
OY 120 NIA 122
DB 137 NIA 139
DR
RESULT 4
DHET_ACEAC STANDARD; PRT; 742 AA.
AC 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Alcohol dehydrogenase [acceptor] precursor (EC 1.1.99.8).
GN ADHA OR ADH1.
OS Acetobacter aceti.
OC Acetobacteria; alpha subdivision; Acetobacteraceae;
OC Acetobacter.
ON NCBI_TaxID=435;
RX RP SEQUENCE FROM N.A., AND SEQUENCE OF 36-50.
RX MEDLINE=8955070; PubMed=2722742;
RA Inoue T., Sunagawa M., Mori A., Imai C., Fukuda M., Takagi M.,
RA Yano K.
RT "Cloning and sequencing of the gene encoding the 72-kilodalton
RT dehydrogenase subunit of alcohol dehydrogenase from Acetobacter
RT aceti."
RL J. Bacteriol. 171:3115-3122(1989).
RL [2]
RP 3D-STRUCTURE MODELING:
RX MEDLINE=95289964; PubMed=7772016;
RX Cozior G.E., Gilles T.G., Anthony C.;
RT "The structure of the quinoprotein alcohol dehydrogenase of
RT Acetobacter aceti modelled on that of methanol dehydrogenase from
RT Methylobacterium extorquens."
RT Biochem. J. 308:375-379(1995).
RL RT
CC -1- CATALYTIC ACTIVITY: A primary alcohol + acceptor = an aldehyde +
CC reduced acceptor.
CC -1- COFACTOR: POQ AND HEME.
CC -1- SUBUNIT: Tetramer of non identical chains (dehydrogenase,
CC cytochrome, and two smaller unknown subunits).
CC -1- SUBCELLULAR LOCATION: Periplasmic.
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POQ DEHYDROGENASE FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; D90004; BAAL14058.1; -.

```

DR	HSP; JS0326;	J50326.
DR	PIR; Q9ZAJ7;	IFLG.
DR	InterPro:	IPR001479; Bac_POO.
DR	InterPro:	IPR002372; Bac_POO_repeat.
DR	InterPro:	IPR000345; CytC_heme_bind.
DR	Pfam:	PF01011; Bacterial_POO_6.
DR	PROSITE:	PS00365; BACTERIAL_POO_1; 1.
DR	PROSITE:	PS00364; BACTERIAL_POO_2; 1.
DR	PROSITE:	PS00190; CYTOCHROME_C; 1.
KW	Oxidoreductase:	POO; Heme; Ferriplasmic; Membrane; Signal.
FT	SIGNAL	1 35
FT	CHAIN	36 742
FT	ACT_SITE	343 742
FT	BINDING	649 649
FT	BINDING	652 652
FT	METAL	653 653
SO	SEQUENCE	742 AA; 81521 MW; 9c6c9268dabb825a CRC64;
Qy	Query Match	26.4%; Score 178; DB 1; Length 742;
Best Local Similarity	37.9%; Pred. No. 7e-10;	
Matches	44; Conservative 19; Mismatches 49; Indels 4; Gaps	
Db	11 AGAALLAAPFAAGTPTVDL--ANPAGEWISYSGNDEYNHSLPLQTITENYGQLD	68
	: : : : :	
23 AALPYAAVVPARADQGNTGATLIHADHDPEMNLSTGRYSSEQRYSLDPINRSNGDLK	82	
Qy	69 LVMARGMOPEGKYQ-VTPLEHGWMYLANPGDIVDAIDAKTGGLIWEHRRLP-NIA	122
	: : : : : : : : : : : : : : : : :	
Db	83 LIIGYTTLDTNRGOEATPLVVGGIMYATTNWSKMEDALDAATGKLTMWDPKPVGNIA	138
RESULT	5	
ID	DHG_ECOLI	STANDARD; PRT; 796 AA.
AC	P15877;	
DT	01-APR-1990 (Rel. 14, Created)	
DT	01-NOV-1997 (Rel. 35, Last sequence update)	
DT	16-OCT-2001 (Rel. 40, Last annotation update)	
DE	Glucose dehydrogenase [pyrroloquinoline-quinone] (EC 1.1.99.17).	
GN	GCD OR B0124.	
OS	Escherichia coli.	
OC	Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;	
OC	Escherichia.	
OX	NCBL_TaxID=562;	
RN	[1]	
RN	SEQUENCE FROM N.A.	
RC	STRAIN-K12;	
RA	MEDLINE-91035240; PubMed-2228962;	
RA	Cleton-Jansen A.-M., Goosen N., Fayet O., van de Putte P.;	
RT	"Cloning, mapping, and sequencing of the gene encoding Escherichia	
RT	coli quinoprotein glucose dehydrogenase.";	
RU	J. Bacteriol. 172:6308-6315(1990).	
RN	[2]	
RN	SEQUENCE FROM N.A.	
RP	STRAIN-K12 / W3110;	
RC	MEDLINE-94261430; PubMed-8419307;	
RA	Yamada M., Asaka S., Saler M.H. Jr., Yamada Y.;	
RT	"Characterization of the gdc gene from Escherichia coli K-12 W3110	
RT	and regulation of its expression";	
RU	J. Bacteriol. 175:568-571(1993).	
RN	[3]	
RN	SEQUENCE FROM N.A.	
RP	STRAIN-K12 / W3110;	
RC	MEDLINE-94261430; PubMed-8202364;	
RA	Fujita N., Mori H., Yura T., Ishihama A.;	
RT	*Systematic sequencing of the Escherichia coli genome: analysis of	
RT	the 2.4.4.1 min (110,917-199,643 bp) region.";	
RU	Nucleic Acids Res. 22:1657-1693(1994).	
RN	[4]	
RN	SEQUENCE FROM N.A.	
RP	STRAIN-K12 / MG1655;	
RC	MEDLINE-97426617; Pubmed-9278503;	
EX		

Query Match	Best Local Similarity	Score 178: DB 1:	Length 742:
Matches 44: Conservative	19: Mismatches	49: Indels	4: Gaps
Qy	11 AGAALLAAPFAOVFTVYTDLLT--ANPAPGEMISYGQNDENYRHSPLQTITENTYGLQ	68	
Db	23 AALPYAALVPRARAGCGNTGATIIHADDPENMLSYGRYSSEQRYSPLDQINRSVGDLEK	82	
Qy	69 LVMARGMOPGKVVQ-VTPLIHDGWTYLANPGDYLQALDAKTKGDLIMFRRLP-NIA	122	
Db	83 ILGYTLTDTNRGQEAFTPLVDGIMYATTNNSKMEALDAATGAKLLMWDPRKVPONIA	138	
RESULT 5			
ID	DHG_ECOLI	STANDARD:	PRT: 796 AA.
AC	P15877:		
DT	01-APR-1990 (Rel. 14, Created)		
DT	01-NOV-1997 (Rel. 35, Last sequence update)		
DT	16-OCT-2001 (Rel. 40, Last annotation update)		
DE	Glucose dehydrogenase [pyrroloquinoline-quinone] (EC 1.1.99.17).		
GN	GCD OR B0124.		
OS	Escherichia coli.		
OC	Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;		
OC	Escherichia.		
OX	NCBI_TaxID=562;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=K12;		
RC	MEDLINE=91035240; PubMed=2228692;		
RA	Cleion-Jansen A.-M., Goosen N., Fayet O., van de Putte P.;		
RT	"Cloning, mapping, and sequencing of the gene encoding Escherichia		
RT	coli quinoprotein glucose dehydrogenase.";		
RL	J. Bacteriol. 172:6308-6315(1990).		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=K12 / W3110;		
RC	MEDLINE=93123180; PubMed=8419307;		
RA	Yamada M., Asosaka S., Seiler M.H. Jr., Yamada Y.;		
RT	"Characterization of the gcd gene from Escherichia coli K-12 W3110		
RT	and regulation of its expression.";		
RL	J. Bacteriol. 175:568-571(1993).		
RN	[3]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=K12 / W3110;		
RC	MEDLINE=94261430; PubMed=8202364;		
RA	Fujita N., Mori H., Yura T., Ishihama A.;		
RT	"Systematic sequencing of the Escherichia coli genome: analysis of		
RT	the 2.4-4.1 min (110.917-193.643 bp) region.";		
RL	Nucleic Acids Res. 22:1637-1639(1994).		
RN	[4]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=K12 / MG1655;		
RX	MEDLINE=97426617; PubMed=9278503;		

RA Blatner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of *Escherichia coli* K-12.";
 RL Science 277:1453-1474(1997).
 RP TOPOLOGY.
 RA MEDLINE=93286127; PubMed=8509415;
 RA Yamada M., Sumi K., Matsushita K., Adachi O., Yamada Y.;
 RT "Topological analysis of quinoprotein glucose dehydrogenase in
 RT *Escherichia coli* and its ubiquinone-binding site.";
 RL J. Biol. Chem. 268:12812-12817(1993).
 RP 3D-STRUCTURE MODELING.
 RX MEDLINE=96128046; PubMed=8554505;
 RX Cozier G.E., Anthony C.;
 RT "Structure of the quinoprotein glucose dehydrogenase of *Escherichia coli* modelled on that of methanol dehydrogenase from *Methylobacterium extorquens*.";
 RL Biochem. J. 312:679-685(1995).
 CC -1- FUNCTION: GDH IS PROBABLY INVOLVED IN ENERGY CONSERVATION RATHER THAN IN SUGAR METABOLISM.
 CC -1- CATALYTIC ACTIVITY: D-glucose + acceptor = D-glucono-1,5-lactone + reduced acceptor.
 CC -1- COFACTOR: POO.
 CC -1- SUBUNIT: MONOMER.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane.
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POO DEHYDROGENASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or_send_an_email_to_license@sib-sib.ch).
 CC -----
 DR EMBL: X51323; CAA35706.1; -;
 DR EMBL: D12651; BAA02174.1; -;
 DR EMBL: D26562; CAB20298.1; -;
 DR EMBL: AE000122; AAC73335.1; -;
 DR PIR: JVO107; JVO107.
 DR HSSP: P38539; 4AAH.
 DR EcoGene: EG10369; gcd.
 DR InterPro: IPR002372; Bac_POO.
 DR Pfam: PF01011; Bacterial_POO_7.
 DR PROSITE: PS00363; BACTERIAL_POO_1; 1.
 DR PROSITE: PS00364; BACTERIAL_POO_2; 1.
 KW Oxidoreductase; POO; Transmembrane; Inner membrane; Periplasmic; Complete proteome.
 FT DOMAIN 1 10
 FT TRANSMEM 11 37
 FT DOMAIN 38 40
 FT TRANSMEM 41 58
 FT DOMAIN 59 62
 FT TRANSMEM 63 81
 FT DOMAIN 82 95
 FT TRANSMEM 96 110
 FT DOMAIN 111 118
 FT TRANSMEM 119 141
 FT DOMAIN 142 796
 FT ACT_SITE 466 466
 FT ACT_SITE 59 59
 FT CONFLICT 149 156
 FT CONFLICT 193 193
 FT CONFLICT 666 666
 SQ SEQUENCE 796 AA; 86747 MW; D9HDC705A12894E9 CRC64;

Query Match 23.0%; Score 155; DB 1; Length 796;
 Best Local Similarity 31.9%; Pred. No. 1.4e-07;

Matches 44; Conservative 20; Mismatches 46; Indels 28; Gaps 6;
 QY 6 LIMS-----AGALALIAAPAFQVTPYDELLANPAGEWISYCONENRHSPLTQ 58
 DB 135 LTMAGFNDPQELNGTSLSDATPAFA-ISPVDQ-----DWAYRNRNGSGRFSFLKQ 185
 QY 59 ITTENYGLQLYMA-----KGMQPGKV--QVTPPLHDGVMYLANPGVIAIDAKTGD 109
 DB 186 INADNVLHKLAWFERGDKVQKPNPDEITNEVTPIKVGDPLYLCTAHORLFLDAASGK 245
 QY 110 LIMEHRRQLPNATLNSF 127
 DB 246 EKMYHD---PELKTNESEF 260
 RESULT 6
 ID EXAA_PSEAE STANDARD; PRT; 623 AA.
 AC G924J7;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Quinoprotein ethanol dehydrogenase precursor (EC 1.1.99.-) (OEDH).
 GN EXAA OR PA1982.
 OS *Pseudomonas aeruginosa*.
 OC Bacteria; Proteobacteria; gamma subdivision; *Pseudomonadaceae*;
 OC *Pseudomonas*.
 OX NCBI_TaxID=287;
 RP [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 17933;
 RX MEDLINE=99041560; PubMed=9826187;
 RA Diehl A., Wintzingerode F., Goerisch H.;
 RA "Quinoprotein ethanol dehydrogenase of *Pseudomonas aeruginosa* is a homodimer: sequence of the gene and deduced structural properties of the enzyme.";
 RL Eur. J. Biochem. 257:409-419(1998).
 RL [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 15692 / PA01;
 RX MEDLINE=20437337; PubMed=10984043;
 RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,
 RA Hickey M.J., Brinkman F.S.L., Huftagle W.O., Kowalik D.J., Lagrou M.,
 RA Gardner R.L., Goltry L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
 RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
 RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
 RA Reiter J., Sater M.H., Hancock R.E.W., Lory S., Olson M.V.;
 RT "Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic pathogen.";
 RL Nature 406:959-964(2000).
 RL [3]
 RP SEQUENCE OF 1-52 FROM N.A.
 RC STRAIN=ATCC 17933;
 RX MEDLINE=99173751; PubMed=10075429;
 RA Schobert M., Goerisch H.;
 RT "Cytochrome c550 is an essential component of the quinoprotein ethanol oxidation system in *Pseudomonas aeruginosa*: cloning and sequencing of the genes encoding cytochrome c550 and an adjacent acetaldehyde dehydrogenase.";
 RL Microbiology 145:471-481(1999).
 RL [4]
 RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF 35-616.
 RC STRAIN=ATCC 17933;
 RX MEDLINE=20202376; PubMed=10736230;
 RA Ketel T., Diehl A., Knaute T., Stezowski J.J., Hoehne W.,
 RA Goerisch H.;
 RT "X-ray structure of the quinoprotein ethanol dehydrogenase from *Pseudomonas aeruginosa*: basis of substrate specificity.";
 RL J. Mol. Biol. 297:961-974(2000).
 CC -1- FUNCTION: OXIDIZE PRIMARY ALCOHOLS AND ALSO ACTS ON SECONDARY ALCOHOL, BUT NOT HIGHLY ACTIVE ON METHANOL.
 CC -1- COFACTOR: POO AND CALCIUM.
 CC -1- SUBUNIT: HOMODIMER.

CC -1- SUBCELLULAR LOCATION: Periplasmic.
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL PQO DEHYDROGENASE FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: A1009858; CAA08896.1; -;
 DR EMBL: AE004624; AAG05370.1; -;
 DR EMBL: AF068264; AAC79657.1; -;
 DR PDB: 1FL6; 30-AUG-00.
 DR InterPro: IPR001479; Bac_PQO.
 DR InterPro: IPR002372; Bac_PQO_repeat.
 DR Pfam: PF01011; Bacterial_PQO_7.
 DR PROSITE: PS00363; BACTERIAL_PQO_1; FALSE_NEG.
 DR PROSITE: PS00364; BACTERIAL_PQO_2; FALSE_NEG.
 KM Oxidoreductase; PQO; Periplasmic; signal; Calcium; 3D-structure;
 KM Complete proteome.
 FT SIGNAL 1 34
 FT CHAIN 35 623 QUINOPROTEIN ETHANOL DEHYDROGENASE.
 FT DISULFID 139 140
 SQ SEQUENCE 623 AA; 68123 MW; 32DDE5DF20B291D6 CRC64;
 Query Match 20.4%; Score 137; DB 1; Length 623;
 Best Local Similarity 31.5%; Pred. No. 6.1e-06;
 Matches 40; Conservative 20; Mismatches 57; Indels 10; Gaps 4;
 QY 3 PLSILMASAGALAL--LAAPAFQVTPVTELLAN--PPAGEMISYGNQENYRHSPLT 57
 DB 9 PAGILRLSLHCLFAVAVALGSAALAKVDWEDIANDDKTGVLYQMGWTHQRMSPK 68
 QY 58 QITENYNGQLQVWARM---QPKVQVPLIHGVMTIANGVDVYQADATGDLIME 113
 DB 69 QVNDVNFKLTLPWMSYSFGDEKORQ-ESQAIYSDGIYVTASYSRFLADARTGRKLT 127
 QY 114 HRRLPLN 120
 DB 128 YNRRLPD 134
 RESULT 7
 DHG_GLUOX STANDARD; PRT; 808 AA.
 AC P27175;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Glucose dehydrogenase [pyrroloquinoline-quinone] precursor
 DE (EC 1.1.99.17).
 GN GDH.
 OS Gluconobacter oxydans (Gluconobacter suboxydans).
 OC Bacteria; Proteobacteria; alpha subdivision; Acetobacteraceae;
 OC Gluconobacter.
 OX NCBI_TaxID=442;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92017653; PubMed=1833618;
 RA Cleton-Jansen A.-M., Dekker S., van de Putte P., Goosen N.;
 RT "A single amino acid substitution changes the substrate specificity
 RT of quinoprotein glucose dehydrogenase in Gluconobacter oxydans.",
 RL Mol. Genet. 229:206-212(1991).
 RN [2]
 RP REVISION TO 213.
 RA Goosen N.;
 RL Submitted (FEB-1992) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: D-glucose + acceptor = D-glucono-1,5-lactone +
 CC reduced acceptor.
 CC -1- COFACTOR: PQO.

CC -1- SUBCELLULAR LOCATION: OUTSIDE OF THE PERIPLASMIC MEMBRANE.
 CC -1- MISCELLANEOUS: THE P1 FORM CAN OXIDIZE ONLY D-GLUCOSE, WHILE THE
 CC P2 FORM CAN ALSO OXIDIZE MALTOSE. THE SEQUENCE OF P1 FORM IS SHOWN
 CC HERE.
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL PQO DEHYDROGENASE FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: X62710; CAA44594.1; ALT_SEQ.
 DR PIR: S17716; QPKEX.
 DR InterPro: IPR001479; Bac_PQO.
 DR InterPro: IPR002372; Bac_PQO_repeat.
 DR Pfam: PF01011; Bacterial_PQO_7.
 DR PROSITE: PS00363; BACTERIAL_PQO_1; 1.
 DR PROSITE: PS00364; BACTERIAL_PQO_2; 1.
 KM Oxidoreductase; PQO; Transmembrane; Periplasmic; signal.
 FT SIGNAL 1 33
 FT CHAIN 34 808 GLUCOSE DEHYDROGENASE [PYRROLOQUINOLINE-
 FT TRANSMEM 35 54 QUINONE].
 FT TRANSMEM 59 76 POTENTIAL.
 FT TRANSMEM 94 108 POTENTIAL.
 FT TRANSMEM 123 138 POTENTIAL.
 FT ACT_SITE 470 470 BASE (POTENTIAL).
 FT VARIANT 788 788 H -> N (IN P2 FORM).
 SQ SEQUENCE 808 AA; 87567 MW; 0F4160DA78652445 CRC64;
 Query Match 19.9%; Score 134; DB 1; Length 808;
 Best Local Similarity 27.8%; Pred. No. 1.6e-05;
 Matches 35; Conservative 20; Mismatches 51; Indels 20; Gaps 3;
 QY 13 ALALLAAPAFQVTPVTELLA-----NPPAGEMISYGNQENYRHSPLTQIT 61
 DB 128 AVLALFASLFTDPHDISGLPQIANSAPADPDNPASEMHAHYGRTQAGDRSPINQINA 187
 QY 62 ENVGOLQVW-----ARGNOPKAV--QVPLIHGVMTIANGVDVYQADATGDLIM 112
 DB 188 TTVSNLKYAMHITKDMNMSNDPGQTEATPIEFNNMTLWCSLHQKFLAVDAGARGNVK 247
 QY 113 EHRRL 118
 DB 248 VYDPKL 253
 RESULT 8
 DHGA_ACTICA STANDARD; PRT; 801 AA.
 AC P03463;
 DT 01-NOV-1988 (Rel. 09, Created)
 DT 01-NOV-1988 (Rel. 09, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Glucose dehydrogenase-A [pyrroloquinoline-quinone] precursor
 DE (EC 1.1.99.17) (Quinoprotein glucose DH) (GDH-A).
 GN GDHA.
 OS Acinetobacter calcoaceticus.
 OC Bacteria; Proteobacteria; gamma subdivision; Moraxellaceae;
 OC Acinetobacter.
 OX NCBI_TaxID=471;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=LMD 79.41;
 RX MEDLINE=88289368; PubMed=3399393;
 RA Cleton-Jansen A.-M., Goosen N., Odle G., van de Putte P.;
 RT "Nucleotide sequence of the gene coding for quinoprotein glucose
 RT dehydrogenase from Acinetobacter calcoaceticus.",
 RL Nucleic Acids Res. 16:6228-6228(1988).

OUTA.ACICA STANDARD; PRT; 809 AA.
 ID OUTA.ACICA 059086;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Quinate/quininate dehydrogenase [pyrroloquinoline-quinone]
 DE (EC 1.1.99.25) (NAD(P)-independent quinate dehydrogenase).
 GN quina.
 OS Acinetobacter calcoaceticus.
 OC Bacteria; Proteobacteria; gamma subdivision; Moraxellaceae;
 OC Acinetobacter.
 OX NCBI_TaxID=471;
 RN [1]
 RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
 RC STRAIN-BD413 / ADP1;
 RX MEDLINE=9505936; PubMed=8002591;
 RA Elsemore D.A., Ornstion L.N.;
 RT "The pca-pob supraoperonic cluster of Acinetobacter calcoaceticus
 RT contains quia, the structural gene for quinate-shikimate
 RT dehydrogenase.";
 RL J. Bacteriol. 176:7659-7666(1994).
 RN [2]
 RP SEQUENCE OF 1-20 FROM N.A.
 RC STRAIN-BD413 / ADP1;
 RX MEDLINE=96011389; PubMed=7592351;
 RA Elsemore D.A., Ornstion L.N.;
 RT "Unusual ancestry of dehydratases associated with quinate catabolism
 RT in Acinetobacter calcoaceticus.";
 RL J. Bacteriol. 177:5971-5978(1995).
 CC -1- FUNCTION: CAN ACT EITHER ON QUINATE OR ON SHIKIMATE.
 CC -1- CATALYTIC ACTIVITY: quinate + pyrroloquinoline-quinone -> 5-
 CC dehydroquininate + reduced pyrroloquinoline-quinone.
 CC -1- CATALYTIC ACTIVITY: SHIKIMATE + PYRROLOQUINOLINE-QUINONE = 3-
 CC DEHYDROSHIKIMATE + REDUCED PYRROLOQUINOLINE-QUINONE.
 CC -1- COFACTOR: POQ.
 CC -1- PATHWAY: QUINIC ACID CATABOLIC PATHWAY. FIRST STEP. THIS PATHWAY
 CC ALLOWS GROWTH OF BACTERIA WITH QUINATE BY ITS CONVERSION TO
 CC PROTOPICACETATE AND SUBSEQUENT METABOLISM BY THE BETA-KETOADIPATE
 CC PATHWAY.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -1- INDUCTION: BY PROTOPICACETATE.
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POQ DEHYDROGENASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: L05770; AAC37161.1; -;
 DR InterPro: IPR001479; Bac_POQ.
 DR Pfam: PF01011; Bacterial_POQ_7.
 DR PROSITE: PS00363; BACTERIAL_POQ_1; FALSE_NEG.
 DR PROSITE: PS00364; BACTERIAL_POQ_2; FALSE_NEG.
 KW Oxidoreductase; POQ; Quinate metabolism; Transmembrane.
 FT TRANSMEM 14 34
 FT TRANSMEM 41 61
 FT TRANSMEM 68 88
 FT TRANSMEM 90 110
 FT TRANSMEM 127 147
 FT TRANSMEM 147 147
 SQ SEQUENCE 809 AA; 88196 MW; 71f67cbea62bfc6 CRC64;

Query Match 14.0%; Score 94.5; DB 1; Length 809;
 Best Local Similarity 29.2%; Pred. No. 0.12;
 Matches 35; Conservative 13; Mismatches 49; Indels 23; Gaps 4;

3 PISLWASGALALLAA-PAFAQVTPVTDLLANPAGMISYGOENYRHSPLTQIT 61

DB 148 PHENVKAGEELPLVPDPARKV-----NMDHGNDAAGSRFVALQINR 193
 OY 62 ENVGQLQIVW-----ARGMQPG-KVQVTPPLIHGVMYLANPGDVIQAIADATGDLIWE 113
 DB 194 NNVSXKLEAMRFRTGDFTTGTGNGCAEDQMPLOVGNKVFCLTPNNIFALDADSGKQLMK 253
 RESULT 13
 XOXF_PARDE STANDARD; PRT; 600 AA.
 ID XOXF_PARDE
 AC P29968;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Putative dehydrogenase XOXF precursor (EC 1.1.99.-).
 GN XOXF.
 OS Paracoccus denitrificans.
 OC Bacteria; Proteobacteria; alpha subdivision; Rhodobacter group;
 OC Paracoccus.
 OX NCBI_TaxID=266;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Harms N.;
 RL Submitted (AUG-1995) to the EMBL/Genbank/DBJ databases.
 RN [2]
 RP SEQUENCE OF 380-600 FROM N.A.
 RC STRAIN-PD 1235;
 RX MEDLINE=92041583; PubMed=1657873;
 RA Ras J., Reijnders W.N.M., van Spanning R.J.M., Harms N., Oltmann L.F.,
 RA Stouthamer A.H.;
 RT "Isolation, sequencing, and mutagenesis of the gene encoding
 RT cytochrome c551 of Paracoccus denitrificans and characterization of
 RT the mutant strain.";
 RL J. Bacteriol. 173:6971-6979(1991).
 CC -1- CORFACTOR: POQ (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POQ DEHYDROGENASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: U34346; AAC44555.1; -;
 DR EMBL: M75583; AAA25574.1; -;
 DR PIR: A41378; A41378.
 DR HSP: P38359; AAAH.
 DR InterPro: IPR002372; Bac_POQ_repeat.
 DR Pfam: PF01011; Bacterial_POQ_7.
 KW Oxidoreductase; POQ; Signal.
 FT SIGNAL 1 21
 FT CHAIN 22 600
 FT ACT_SITE 318 318
 FT ACT_SITE 318 318
 SQ SEQUENCE 600 AA; 65159 MW; DCA9396f1bcca3ce CRC64;

Query Match 13.8%; Score 93; DB 1; Length 600;
 Best Local Similarity 27.9%; Pred. No. 0.12;
 Matches 36; Conservative 23; Mismatches 48; Indels 22; Gaps 7;

10 SAGATALLAFAFAQVTPVTDLLANP-----PAGEWISYGOENYRHSPLTQITTE 62
 DB 6 NCACIALALMSGIAA-----LANEORAGRDRQAQPMALQMGDIYANTRISTLDQINKD 56

63 NVGQLQIVW--ARGMQPGKVQVTPPLIHGVMYLANP-GDVIOAIDAK-TGDLIWEHR-RQ 117
 DB 57 NVKDLRVAMFTSTGVLRGH-EGSPVIGDVYVYHTPPNNRVFALDLNDCKIILRYEPQ 115

118 LPNATLNS 126
 DB 116 DPNVIAVMS 124

Fri May 24 11:27:17 2002

us-08-934-506a-5_copy_1_128.rsp

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2002, 10:17:30 ; Search time 93.93 Seconds
(without alignments)
235.743 Million cell updates/sec

Title: US-08-934-506A-5_COPY_1_128
Sequence: 1 MKPTSLMAGALALAAAP.....DLIWEHRQLPNIATLNSFG 128

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_19:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_proteome:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	261	38.8	608	2	093RE9
2	190	28.2	742	2	053362
3	178.5	26.5	691	2	09AF95
4	174	25.9	623	2	09AGW3
5	168.5	25.0	698	2	09KH03
6	167	24.8	786	2	P95466
7	165.5	24.6	708	2	046444
8	156	23.2	695	2	09F9U2
9	155	23.0	777	16	09RKF6
10	146.5	21.8	777	16	09RFB3
11	137.5	20.4	803	16	091115
12	134.5	20.0	601	2	09EYW8
13	134.5	20.0	601	16	092MY9
14	121	18.0	790	2	09XZS5
15	118.5	17.6	695	2	0934G0
16	112	16.6	601	2	P71509

17	110.5	16.4	629	2	09A048	09ag48 methylolact
18	103	15.3	633	2	024759	024759 hypomicrob
19	93.5	13.9	639	2	P77931	P77931 pseudomonas
20	93.5	13.9	644	2	052551	052551 pseudomonas
21	89	13.2	499	2	051842	051842 thiolactillu
22	83.5	12.4	2832	2	093N05	093N05 agrobacteri
23	83	12.3	599	2	09L935	09L935 methylovoru
24	82.5	12.3	485	17	096XP7	096XP7 sulfolobus
25	82	12.2	668	17	09RRH3	09RRH3 deinococcus
26	81	12.0	613	17	028873	028873 archaeoglob
27	80.5	12.0	647	16	09HDB1	09HDB1 pseudomonas
28	78.5	11.7	509	16	P6368	P6368 mycobacteri
29	77	11.4	846	2	086092	086092 methylovoru
30	76	11.3	306	16	09R772	09R772 deinococcus
31	76	11.3	386	16	09RTW8	09RTW8 vibrio chol
32	76	11.3	573	2	059540	059540 methylophil
33	75.5	11.2	407	16	09PG38	09PG38 xylella fas
34	75.5	11.2	684	2	09S282	09S282 streptomyce
35	75.5	11.2	3972	16	P73139	P73139 synecocyst
36	75	11.1	680	5	09N9G7	09N9G7 phallusia m
37	74	11.0	380	16	09HXJ7	09HXJ7 pseudomonas
38	74	11.0	525	16	0931E9	0931E9 staphylococ
39	74	11.0	755	2	09MNA1	09MNA1 agrobacteri
40	74	11.0	969	16	09RZ15	09RZ15 deinococcus
41	73.5	10.9	183	16	09CKL1	09CKL1 pasteurella
42	73.5	10.9	1229	5	09VW05	09VW05 drosophila
43	73	10.8	355	16	0926F4	0926F4 rhizobium m
44	73	10.8	407	16	09A6E2	09A6E2 caulobacter
45	73	10.8	439	10	004099	004099 brassica ca

ALIGNMENTS

RESULT 1
ID 093RE9 PRELIMINARY; PRT; 608 AA.
AC 093RE9
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE ALCOHOL DEHYDROGENASE.
GN ADH.
OS Pseudoglucobacter saccharotogenes.
OC Bacteria; Pseudoglucobacter.
OX NCBI_TaxID=133921;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IFO 14464;
RA Shidara T., Saito Y.;
RT "Alcohol dehydrogenase.";
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB046580; BAB62258.1;
SQ SEQUENCE 608 AA; 65101 MW; 0ACCE97AE1BA570 CRC64;

Query Match 38.8%; Score 261; DB 2; Length 608;

Best Local Similarity 41.0%; Pred. No. 1.6e-17;
Matches 57; Conservative 20; Mismatches 42; Indels 20; Gaps 2;

QY 10 SAGALALAAPFAO-----VTPTDELANPAGEWISYGNQEN 50
DB 16 STALLASLSGPAFAODHNAAPSPKAGOSALIENTOPTYADDLACKNPNMPLTNGNG 75
QY 51 YHSPPLTQITTEVNGOLQVWARGMQPKGVQYTPLIHGVWYLANPGVYQAIIDAKTGL 110
DB 76 WGSYPLDQINKDNVGLQVWSTRTMPEGSNEGAALAYNVIFLGNTVIAIDKGTSL 135
QY 111 IWEHRQLPNIATLNSFG 128
DB 136 IWEYRRKLPASAKFINSIG 154

RESULT 2

ID 053362 PRELIMINARY; PRT; 742 AA.

AC 053362; 044159; 01, Created)

DT 01-NOV-1996 (TREMBLrel. 01, last sequence update)

DE 01-OCT-2001 (TREMBLrel. 18, last annotation update)

OS Acetobacter pasteurianus (Acetobacter turbianus).

OC Bacteria; Proteobacteria; alpha subdivision; Acetobacteraceae;

NCBI_TaxID=438;

RP SEQUENCE FROM N.A.

RC STRAIN-NC11380;

RX MEDLINE-94042848; PubMed-8226628;

RA Takemura H., Kondo K., Horiuchi S., Beppu T.;

RT Induction by ethanol of alcohol dehydrogenase activity in Acetobacter pasteurianus.

RL J. Bacteriol. 175:6857-6866(1993).

DR EMBL; D13893; BAA40252.1; -.

DR HSSP; 0924J7; 1FLG.

DR InterPro; IPR001479; Bac_PQO.

DR InterPro; IPR002372; Bac_PQO_repeat.

DR Pfam; PF01011; Bacterial_PQO; 6.

DR PROSITE; PS00363; BACTERIAL_PQO_1; 1.

DR PROSITE; PS00364; BACTERIAL_PQO_2; 1.

SEQUENCE 742 AA; 81680 MW; 031C4E9B6B13EC98 CRC64;

Query Match 28.2%; Score 190; DB 2; Length 742;

Best Local Similarity 38.8%; Pred. No. 2.2e-10;

Matches 45; Conservative 18; Mismatches 49; Indels 4; Gaps 3;

QY 11 AGALLIAPAPAOYTPYDEL--ANPPAGEMISYQNGQENYRHSPLQTITTEWGLQ 68

DB 23 AALPAPAPADGCGTNGTGTADHDPENMLSTGYSPQSPDQINRSWGLK 82

QY 69 LWARGMOPGKVO-VTPLIHDGVMTLANPGDVIQAIDAKTGDLWEHRRQIP-NIA 122

DB 83 LAMVYTLDTNRGQENATPLVVDGIMVATNWSKMLADATGKLMQYDPKVPGNIA 138

RESULT 3

ID 09AF95 PRELIMINARY; PRT; 691 AA.

AC 09AF95;

DT 01-JUN-2001 (TREMBLrel. 17, Created)

DT 01-DEC-2001 (TREMBLrel. 19, last sequence update)

DE 01-DEC-2001 (TREMBLrel. 19, last annotation update)

OS Pseudomonas butanovora.

OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;

NCBI_TaxID=86174;

RP SEQUENCE FROM N.A.

RA Vangnai A.S., Arp D.J., Sayavedra-Soto L.A.;

RT "Characterization of the expression of two distinct alcohol dehydrogenases involved in butane metabolism in Pseudomonas butanovora."

RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF35798; AKR72220.2; -.

SEQUENCE 691 AA; 75070 MW; 4FC7FD20CDA14E64 CRC64;

Query Match 26.5%; Score 178.5; DB 2; Length 691;

Best Local Similarity 37.8%; Pred. No. 2.8e-09;

Matches 31; Conservative 22; Mismatches 28; Indels 1; Gaps 1;

QY 39 GEMWISYQNGQENYRHSPLQTITTEWGLQDLWARGMOPGK-VQVTPLIHDGVMTLANPG 97

DB 41 GEMWTHGVDAAGTRYSPLAQITTPNNAKELGLWVSYDLESSRGVATITVVDGVMYVIAFW 100

QY 98 DVIAIDAKTGDLWEHRRQIP 119

DB 101 SVVHALDVRSKRLMTYDPEVP 122

RESULT 4

ID 09AGW3 PRELIMINARY; PRT; 623 AA.

AC 09AGW3;

DT 01-JUN-2001 (TREMBLrel. 17, Created)

DT 01-JUN-2001 (TREMBLrel. 17, last sequence update)

DE 01-DEC-2001 (TREMBLrel. 19, last annotation update)

OS Pseudomonas butanovora.

OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;

NCBI_TaxID=66174;

RP SEQUENCE FROM N.A.

RA Sayavedra-Soto L.A., Vangnai A.S., Arp D.J.;

RT "Characterization of the expression of two distinct alcohol dehydrogenases involved in butane metabolism in Pseudomonas butanovora."

RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF326086; AAK15506.1; -.

DR HSSP; 0924J7; 1FLG.

DR InterPro; IPR002372; Bac_PQO_repeat.

DR Pfam; PF01011; Bacterial_PQO; 4.

FW Signal.

FT SIGNAL.

SEQUENCE 623 AA; 67553 MW; A573A1C64AC12D55 CRC64;

Query Match 25.9%; Score 174; DB 2; Length 623;

Best Local Similarity 38.0%; Pred. No. 6.8e-09;

Matches 49; Conservative 18; Mismatches 46; Indels 16; Gaps 5;

QY 2 KPTSLWASGALIALAPFAOYTPYDEL--PPAGEMISYQNGQENYRHSPLQI 59

DB 7 KPRALRAIVVATNAUSLPA-AAVDTWEDINDHRTGDTLYLGAKQHSPLKAI 65

QY 60 TTEWGLQDLWMA-----RGMOPGKVOVTPLIHDGVMTLANPGDVIQAIDAKTGDL 111

DB 66 NTIDVNAVLPAMSPFSFGKEKQGE---AGV--LVHDVYIYTAAS.SRIFAIDARSGRL 120

QY 112 WEHRRQIPN 120

DB 121 WEYNARLPD 129

RESULT 5

ID 09KH03 PRELIMINARY; PRT; 698 AA.

AC 09KH03;

DT 01-OCT-2000 (TREMBLrel. 15, Created)

DT 01-OCT-2000 (TREMBLrel. 15, last sequence update)

DE 01-DEC-2001 (TREMBLrel. 19, last annotation update)

OS Alcaligenes eutrophus (Ralstonia eutropha).

OC Bacteria; Proteobacteria; beta subdivision; Ralstonia group;

NCBI_TaxID=510;

RP SEQUENCE FROM N.A.

RX MEDLINE-21125557; PubMed-11222593;

RA Zarrin G., Schrader T., Andreesen J.R.;

RT "Catalytic and Molecular Properties of the Quinolomoprotein Tetrahydrofururyl Alcohol Dehydrogenase from Ralstonia eutropha Strain Bo."

RL J. Bacteriol. 183:1954-1960(2001).

DR EMBL; AF277373; AAF6335.1; -.

HSSP; 0924J7; 1FLG.

DR Interpro: IPR002375: Bac_PQO_repeat.
DR Interpro: IPR000345: CYC_heme_bind.
DR Interpro: IPR003088: Cyt_Ct.
DR Pfam: PF01011: Bacterial_POQ_6.
DR Pfam: PF00034: cytochrome_c_1.
DR PROSITE: PS00190: CYTOCHROME_C; UNKNOWN_1.
SQ SEQUENCE 698 AA: 7541 MW: 963718DBD1DFB2 CRC64:

Query Match	25.0%	Score 168.5;	DB 2;	Length 698;
Best Local Similarity	36.5%	Pred. No. 2.8e+08;		
Matches 42;	Conservative 19;	Mismatches 49;	Indels 5;	Gaps 3;

```

OY  SAGLALILAMPAPA--QVYPIVDELLAMPAG--EMISYCONCENYHNSPLTITENNG 65
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
14  AASVVALPAPMARGAANAARVOCALRANENAGTPWMPSTYGLDYAEYRFSKLEQVNGNVR 73
Db
OY
Db  66  QOLQVWARGQOPK--QVYPIILHDCVWLANGVOLIAIDAKTDDLWEHRRLDP 119
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
74  NUGLAMSVIDTDSRGVEARPLVVDGEMAYISAPWASVVALIDAKTRKMTYDPOVF 128

```

RESULT	6
P95466	
ID	P95466
PRELIMINARY:	
PRT:	786 AA

DF	01-MAY-1997 (TREMblrel. 03, Created)
DF	01-MAY-1997 (TREMblrel. 03, Last sequence update)
DT	01-JUN-2001 (TREMblrel. 17, last annotation update)
DE	GLUCOSE DEHYDROGENASE [PYRROLOQUINOLINE-QUINONE] (EC 1.1.99.17).
GN	Gdh.
OS	Pantoea citrea.
OC	Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC	Pantoeae.
OX	NCBI_TaxID=53336;
RA	[1]
RP	SEQUENCE FROM N.A., AND ACTIVITY.
RC	STRAIN=1056r;
RX	MEDLINE=97133947; PubMed=8979341;
RA	Cha J.-S., Pujol C., Kado C.I.;
RT	"Identification and characterization of a Pantoea citrea gene encoding
RT	glucose dehydrogenase that is essential for causing pink disease of
RT	pineapple.";
RL	Appl. Environ. Microbiol. 63:771-76(1997).
CC	-1- CATALYTIC ACTIVITY: D-GLUCOSE + ACCEPTOR = D-GLUCONO-1,5-LACTONE +
CC	REDUCED ACCEPTOR.
CC	-1- COFACTOR: POO.
CC	-1- SUBCELLULAR LOCATION: PERIPLASMIC (BY SIMILARITY).
CC	-1- MISCELLANEOUS: RESPONSIBLE FOR PINK TO BROWN DISCOLORATIONS OF
CC	INFECTED PINEAPPLES UPON HEATING DURING THE CANNING PROCEDURE.
CC	INFECTED BATCHES HAVE TO BE DISCARDED.
CC	-1- SIMILARITY: BELONGS TO THE BACTERIAL POO DEHYDROGENASE FAMILY.
DR	EMBL; X95985; CA65229.1; -.
DR	Interp: IPR001479; Bac.POO.
DR	Interp: IPR002372; Bac.POO.repeat.
DR	Pfam: PF01011; Bacterial_POO. 6.
DR	PROSITE; PS00364; BACTERIAL_POO_2; 1.
KW	Oxidoreductase; POO; Transmembrane; periplasmic.
FT	DOMAIN 1 13 CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM 14 55 POTENTIAL.
FT	DOMAIN 56 59 PERIPLASMIC (POTENTIAL).
FT	TRANSMEM 60 80 POTENTIAL.
FT	DOMAIN 81 85 CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM 86 106 POTENTIAL.
FT	DOMAIN 107 115 PERIPLASMIC (POTENTIAL).
FT	TRANSMEM 116 136 POTENTIAL.
FT	DOMAIN 137 431 CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM 432 452 POTENTIAL.
FT	DOMAIN 453 652 PERIPLASMIC (POTENTIAL).
FT	TRANSMEM 653 673 POTENTIAL.
FT	DOMAIN 674 700 CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM 701 721 POTENTIAL.
FT	DOMAIN 722 786 PERIPLASMIC (POTENTIAL).

FT	DOMAIN	724	770	PQQ-BINDING, BY SIMILARITY.
FT	ACT_SITE	462	462	
SQ	SEQUENCE	786 AA;	86038 MW;	372402AAD8B067CC CRC64,

Query Match	24.8%;	Score 167;	DB 2;	Length 786;
Best Local Similarity	33.3%;	Pred. No. 4.5e-08;		
Matches 48;	Conservative 19;	Mismatches 51;	Indels 26;	Gaps 5

```
QY      2 KPTSLTMSACALALLAPAPAQ-----VTPVTDELANP-----AGWISYGQNG 48
        || : ||| |
Db      115 KPLTAMIGIINALLLGASLHPDQINGLVSDK----PPAESAAASLAADPAIRFTQ 170
        || : ||| |
QY      49 ENYRHSPLTOTTENVGOLQJLV-----ARGMQPKV--OYPLIHDSVMYLANGDV 99
        || : ||||| |
Db      171 EGVRSPLTLQINDKNVOQLVAMQFHTGDHKTANDPGEITNEYTPLKVGNNLYCTPHOI 230
        || : ||||| |
QY      100 IQAIDAKTGDLIEWHRQLPINAT 123
        : : || : |
Db      231 LIALDAAGREKWRPDLPQKSDPT 254
```

RESULT	7	
Q46444		
ID	Q46444	PRELIMINARY;
		PRT; 708 AA

DT 01-NOV-1998 (TREMBLrel. 08, Created)
 DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE QUTINOHAEMOPROTEIN ETHANOL DEHYDROGENASE TYPE I PRECURSOR
 DE (EC 1.1.99.-) (QH-BDHL).
 OHBDH.
 GN Comamonas testosteroni (Pseudomonas testosteroni).
 OC Bacteria; Proteobacteria; beta subdivision; Comamonadaceae; Comamonas
 OC NCBI_TaxID=285;
 RN [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN-ATCC 15667;
 RX MEDLINE=96184549; PubMed=8654419;
 RA Stoorvogel J., Kraayveld D.E., van Sluis C.A., Jongejan J.A.,
 RA De Vries S., Duine J.A.;
 RT "Characterization of the gene encoding quinohaemoprotein ethanol
 RT dehydrogenase of Comamonas testosteroni.";
 RL Eur. J. Biochem. 235:690-698(1996).
 RN [2]
 RN SEQUENCE OF 32-54 AND 477-490.
 RP STRAIN-ATCC 15667;
 RC MEDLINE=95324580; PubMed=7601151;
 RX De Jong G.A.H., Geerlof A., Stoorvogel J., Jongejan J.A., De Vries S.,
 RA Duine J.A.;
 RT "Quinohaemoprotein ethanol dehydrogenase from Comamonas testosteroni.
 RT Purification, characterization, and reconstitution of the apoenzyme
 RT with pyrioloquinoline quinone analogues.";
 RL Eur. J. Biochem. 230:899-905(1996).
 RN [3]
 RN CHARACTERIZATION.
 RP MEDLINE=86242113; PubMed=3521592;
 RX Groen B.W., van Kleef M.A., Duine J.A.;
 RA "Quinohaemoprotein alcohol dehydrogenase apoenzyme from Pseudomonas
 RT testosteroni.";
 RL Biochem. J. 234:611-615(1986).
 RN [4]
 RN CRYSTALLIZATION.
 RP MEDLINE=21536088; PubMed=11679760;
 RX Duine J.A., Hutzinger E.G., Kozeeboom H.J., Kalk K.H., de Jong G.A.H.,
 RA Dulle A., Dijkstra B.W.;
 RT "Crystallization of quinohaemoprotein alcohol dehydrogenase from
 RT Comamonas testosteroni: crystals with unique optical properties.";
 RL Acta Crystallogr. D 57:1732-1734(2001).
 CC -I- FUNCTION: CATALYZES THE DYCE-LINKED OXIDATION OF PRIMARY ALCOHOLS
 CC TO THE CORRESPONDING ALDEHYDES AND THE (SUBSEQUENT) OXIDATION OF
 CC THE ALDEHYDES TO CARBOXYLIC ACIDS.
 CC -I- COFACTOR: POQ, HEME, AND CALCIUM.

CC -1- SUBUNIT: MONOMER.
 CC -1- SUBCELLULAR LOCATION: PERIPLASMIC (POTENTIAL).
 CC -1- INDUCTION: BY ETHANOL AND BUTANOL.
 CC -1- SIMILARITY: TO ALCOHOL DEHYDROGENASE (ACCEPTOR) PROTEINS.
 DR EMBL: X81880; CAA57464.1; -.
 DR HSP: Q9Z4J7; IFLG.
 DR InterPro: IPR002372; Bac_PQQ_repeat.
 DR InterPro: IPR003088; Cyt_C1.
 DR InterPro: IPR002329; Cyt_C1C.
 DR Pfam: PF01011; Bacterial_PQQ; 6.
 DR Pfam: PF00034; cytochrome_c; 1.
 DR PRINTS: PR00605; CYTOCHROME_C; 1.
 KW Signal; PQQ; Heme; Calcium; Oxidoreductase; Periplasmic.
 FT SIGNAL 1 31
 FT CHAIN 32 708 QUINOAEMOPROTEIN ETHANOL DEHYDROGENASE
 FT TYPE 1.
 FT BINDING 635 635 HEME (COVALENT) (BY SIMILARITY).
 FT BINDING 638 638 HEME (COVALENT) (BY SIMILARITY).
 FT METAL 639 639 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
 SQ SEQUENCE 708 AA; 76822 MW; 99AB54BDD6ACAB3 CRC64;

Query Match
 Best Local Similarity 24.6%; Score 165.5; DB 2; Length 708;
 Matches 41; Conservative 23; Mismatches 50; Indels 15; Gaps 4;

QY 3 PTSLMASGALALLAFAFAQVTPVT-----DELLANPP-AGENISYQONEN 50
 DB 12 PERWVWLAAAGC--SAAFATGTGAQAQAAAVRVDDGFITANAARPPDMFTIGVDAE 69
 QY 51 YHSHPLTQTTENVGQLOLVMAKQPGK-VQVPLIHGVYTLANPGDVIQAIDAKTGD 109
 DB 70 TYSRSDQINANVNDLGLAMSYNLESTREGEATPVVDGIMVYSASVVAIDITRGN 129
 QY 110 LIEHRRPOT 118
 DB 130 RIMTYDPQI 138

RESULT 8
 Q9F902 PRELIMINARY; PRT; 695 AA.
 AC Q9F902:
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, last annotation update)
 DE ALCOHOL DEHYDROGENASE.
 GN ADH.
 OS Pseudomonas stutzeri (Pseudomonas perfectomarina).
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 OC Pseudomonas.
 OX NCBI_TaxID=316;
 OX [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=BC-2;
 RA Chang C.-H., Herrick J.B., Okinaka R.T., Brainard J.B.,
 RA "Tewilliger T.C.;
 RT "Identification and characterization of genes activated by 2-
 RT chloroethanol in Pseudomonas stutzeri BC-2.";
 RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF176640; AAC09249.1; -.
 DR HSP: Q9Z4J7; IFLG.
 DR InterPro: IPR002372; Bac_PQQ_repeat.
 DR InterPro: IPR003088; Cyt_C1.
 DR InterPro: IPR002329; Cyt_C1C.
 DR Pfam: PF01011; Bacterial_PQQ; 6.
 DR Pfam: PF00034; cytochrome_c; 1.
 DR PROSITE, PS00190; CYTOCHROME_C; UNKNOWN_1.
 SQ SEQUENCE 695 AA; 75842 MW; D29A698A642E1A83 CRC64;

Query Match
 Best Local Similarity 23.2%; Score 156; DB 2; Length 695;
 Matches 33.6%; Pred. No. 4.8e-07;

Matches 36; Conservative 22; Mismatches 45; Indels 4; Gaps 3;
 QY 15 ALLAAPAFQVTPVTDLL--ANPPAGEWISYQONENRHSPLTQTTENVGQLOLVMA 72
 DB 16 ALLVA-AGAAQAKYDEAIRASEODGSEWLSHGRTVAEQRSPKQIDAGNVGRLGAWY 74
 QY 73 RGMOPGK-VQVPLIHGVYTLANPGDVIQAIDAKTGDLLWEHROL 118
 DB 75 LDLENKGLKATPLVSDGLVSLASMSRWMAVDLRSGKRLMQEDPOV 121

RESULT 9
 Q98KF6 PRELIMINARY; PRT; 785 AA.
 AC Q98KF6:
 DT 01-OCT-2001 (TREMBLrel. 18, Created)
 DT 01-OCT-2001 (TREMBLrel. 18, last sequence update)
 DT 01-OCT-2001 (TREMBLrel. 18, last annotation update)
 DE GLUCOSE DEHYDROGENASE.
 GN ML1500.
 OS Rhizobium loti (Mesorhizobium loti).
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Phyllobacteriaceae; Mesorhizobium.
 OX NCBI_TaxID=381;
 OX [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=MAF303099;
 RX MEDLINE=21082930; PubMed=11214968;
 RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
 RA Matanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
 RA Kishida Y., Kiyokawa C., Kohara M., Matsunoto M., Matsuno A.,
 RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
 RA "Teuchi C., Yamada M., Tabata S.;
 RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
 RT Mesorhizobium loti.";
 RL DNA Res. 7:331-338(2000).
 DR EMBL: AP002997; BAB48858.1; -.
 DR InterPro: IPR002372; Bac_PQQ_repeat.
 DR Pfam: PF01011; Bacterial_PQQ; 7.
 KW Complete proteome.
 SQ SEQUENCE 785 AA; 84083 MW; 1DC224B7A1C25470 CRC64;

Query Match
 Best Local Similarity 23.0%; Score 155; DB 16; Length 785;
 Matches 46; Conservative 13; Mismatches 46; Indels 40; Gaps 4;

QY 8 WASAG-----ALALLAPAFQVTPVTD-----ELLANPAGE 40
 DB 111 WARRGLAGDPGRAPLLIAYLASIYAVGSMETADPKIGDALDPTDKVTPKANGLNDVPAGE 170
 QY 41 WISYQONENRHSPLTQTTENVGQLOLVMAKQPGKV-----QVTPLIHNG 89
 DB 171 WHYGRTOFGQSRPSLDQITTPDVNAMLQPAWT--YRTGVKGPDDIGETTYVTPPLKIGD 228
 QY 90 VMTIANGDVITQAIIDAKTGDLLWEH 114
 DB 229 TLVTCRPHNPALAVDAATCKEKRRY 253

RESULT 10
 Q92RB3 PRELIMINARY; PRT; 777 AA.
 AC Q92RB3:
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)
 DE PROBABLE GLUCOSE DEHYDROGENASE (PYRROLOQUINOLINE-QUINONE) PROTEIN
 DE (EC 1.1.99.17).
 OS Rhizobium meliloti (Sinorhizobium meliloti).
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Rhizobiaceae; Sinorhizobium.
 OX NCBI_TaxID=382;

Fri May 24 11:27:18 2002

us-08-934-506a-5_copy_1_128.rspt

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2002, 10:05:01 ; Search time 117.28 Seconds

(without alignments)
127.856 Million cell updates/sec

Title: US-08-934-506a-5_COPY_1_135

Perfect score: 710
Sequence: 1 MKPSTLLMSAGALLLAAP.....ROLPNITATLNFSEPTRGMA 135

Scoring table: BLOSUM62
Gapop 10.0 , Gapept 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

A_Geneseq_032802.*
1: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1980.DAT.*
2: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1981.DAT.*
3: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1982.DAT.*
4: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1983.DAT.*
5: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1984.DAT.*
6: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1985.DAT.*
7: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1986.DAT.*
8: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1987.DAT.*
9: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1988.DAT.*
10: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1989.DAT.*
11: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1990.DAT.*
12: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1991.DAT.*
13: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1992.DAT.*
14: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1993.DAT.*
15: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1994.DAT.*
16: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1995.DAT.*
17: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1996.DAT.*
18: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1997.DAT.*
19: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1998.DAT.*
20: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA1999.DAT.*
21: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA2000.DAT.*
22: /SIDSL/gcgdata/hold-geneseq/geneeqp-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	710	100.0	579	19 AAM37873	Alcohol and/or ald
2	640	90.1	579	19 AAM37874	Alcohol and/or ald
3	587.5	82.7	578	19 AAM37875	Alcohol and/or ald
4	567	79.9	579	19 AAM37876	Alcohol and/or ald
5	193.5	27.3	754	21 AAB35987	Sorbitol dehydrog
6	187	26.3	738	12 AAR13993	A.alcoetigines me
7	187	26.3	738	13 AAR20192	ADH complex protei
8	178	25.1	742	11 AAR05235	Amino acid sequenc
9	158.5	22.3	443	22 ABG24430	Novel human diagno
10	145.5	20.5	740	20 AAM95019	Sorbitol dehydrog
11	129	18.2	948	22 ABG25909	Novel human diagno

12	129	18.2	1510	22 ABG21573	Novel human diagno
13	77.5	10.9	353	22 ABA47447	MOAP P5. Haemophi
14	76	10.7	824	18 AAM09614	Pathogenic Staphyl
15	75.5	10.6	1419	22 ABB59208	Drosophila melanog
16	74	10.4	282	22 AAB81927	Acromonium cellulo
17	74	10.4	294	21 AAB35811	Chitinase amino ac
18	74	10.4	294	21 AAB13001	Chitinase protein
19	73.5	10.4	1229	22 ABB71736	Drosophila melanog
20	72.5	10.2	343	19 AAM055001	Bacillus thermoloe
21	72.5	10.2	462	17 AAM05395	Human SH3P17 prote
22	71	10.0	526	22 AAB99895	Xanthomonas sp. ch
23	70.5	9.9	171	22 AAB32330	Novel human secret
24	70.5	9.9	475	22 AAG92298	C glutamicum prote
25	70.5	9.9	843	20 AAI34485	Porphyromonas ging
26	70.5	9.9	849	20 AAI34360	Porphyromonas ging
27	70.5	9.9	5435	22 AAE10145	Streptomyces nous
28	70	9.9	1016	22 AAI34349	Staphylococcus aur
29	69.5	9.8	275	21 AAG42781	Arabidopsis thalia
30	69.5	9.8	1029	22 ABB26194	Novel human diagno
31	69.5	9.8	1029	22 ABB28658	Novel human diagno
32	69	9.7	711	22 ABB69743	Drosophila melanog
33	69	9.7	957	22 ABB62401	Drosophila melanog
34	68.5	9.6	384	22 ABB85255	Plant porpobilling
35	68.5	9.6	551	22 AAB85251	Plant thioedoxin-
36	68.5	9.6	551	22 AAB85252	Plant thioedoxin-
37	68	9.6	159	19 AAM63034	Mycobacterium tube
38	68	9.6	159	21 AAB14325	Mycobacterium tube
39	68	9.6	424	22 ABB59716	Drosophila melanog
40	68	9.6	718	22 AAG91907	C glutamicum prote
41	67.5	9.5	275	21 AAG16656	Arabidopsis thalia
42	67.5	9.5	735	22 ABB24896	Novel human diagno
43	67.5	9.5	851	14 AAR41333	113 kd ISGF-3alpha
44	67.5	9.5	851	16 AAR72077	Recognition factor
45	67.5	9.5	851	17 AAM03166	Human STAT2. Homo

ALIGNMENTS

RESULT 1	
AAW37873	AAW37873 standard; Protein: 579 AA.
ID	AAW37873
XX	10-AUG-1998 (first entry)
AC	AAW37873;
XX	
DT	10-AUG-1998 (first entry)
XX	
DE	Alcohol and/or aldehyde dehydrogenase A amino acid sequence.
XX	
KW	Alcohol/aldehyde dehydrogenase A enzyme; recombinant organism;
KW	aldehyde; ketone; carboxylic acid; L-sorbose; D-sorbitol;
XX	2-keto-L-gulonate acid; L-ascorbic; inhibition.
XX	
OS	Glucanobacter oxydans.
XX	
FH	Key
FT	Peptide
FT	Location/Qualifiers
FT	/note="signal peptide"
FT	24..579
FT	/note="mature protein"
XX	
PN	EP832974-A2.
XX	
PD	01-APR-1998.
XX	
PF	11-SEP-1997; 97EP-0115801.
XX	
PR	19-SEP-1996; 96EP-0115001.
XX	
PA	(HOEF) HOFFMANN LA ROCHE & CO AG F.
XX	
PI	Asakura A, Hoshino T, Ojima S, Shinjoh M, Tomiyama N;
XX	

DR WPI: 1998-195228/18.
DR N-PSDB: AAV29051.
XX
PT Recombinant Gluconobacter oxydans alcohol and/or aldehyde
PT dehydrogenase enzyme(s) - useful for converting substrate(s), e.g.
PT L-sorbose or D-sorbitol to 2-keto-L-gulonic acid
XX
PS Claim 1: Pages 35-37; 59pp; English.
XX
CC This is the amino acid sequence for the Gluconobacter oxydans
CC alcohol and/or aldehyde dehydrogenase A' enzyme. The enzymes
CC or recombinant organisms can be used to convert suitable substrates
CC to aldehydes, ketones or carboxylic acids, especially to convert
CC L-sorbose or D-sorbitol to 2-keto-L-gulonic acid, which can be
CC converted to L-ascorbic acid by standard procedures. The derivatives
CC of ADH enzymes have desired substrate specificity, higher affinity
CC to a substrate, lower affinity to an inhibitory compound, higher
CC stability against temperature and/or pH and higher catalytic speed.
XX
SQ Sequence 579 AA:

Query Match 100.0%; Score 710; DB 19; Length 579;
Best Local Similarity 100.0%; Pred. No. 1,1e-69;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKPTSLMASAGALIALAPAFAYPTVDLLANPAGENISYSGONENYRHSPLTQIT 60
DB 1 mpptslmasagalaialapafayptvdellanpagsewisysqnenyrrhsplqit 60
QY 61 TENVGQLQVWARGMKGQKQVPTPLIHGVMYLANPGDVYQAIDAKTGDLIWEHRRQLPN 120
DB 61 tenvgqlqvwargmkgkqvptplihgvmylanpgdvyaiaidaktgdliwehrrqlpn 120
QY 121 IATINSFGPEPTRGMA 135
DB 121 iatinsfgeptrgma 135

RESULT 2

AAW37874
ID AAW37874 standard; Protein: 579 AA.

XX AAW37874;

DT 10-AUG-1998 (first entry)

DE Alcohol and/or aldehyde dehydrogenase A' amino acid sequence.

XX Alcohol/aldehyde dehydrogenase A' enzyme; recombinant organism;

KW aldehyde; ketone; carboxylic acid; L-sorbose; D-sorbitol;

XX 2-keto-L-gulonic acid; L-ascorbic; inhibition.

OS Gluconobacter oxydans.

XX Key Location/Qualifiers

FT Peptide 1..23

FT Protein /note= "signal peptide" 24..579

FT /note= "mature protein"

XX EP832974-A2.

XX 01-APR-1998.

XX 11-SEP-1997; 97EP-0115801.

XX 19-SEP-1996; 96EP-0115001.

XX (HOFF) HOFFMANN LA ROCHE & CO AG F.

XX Asakura A, Hoshino T, Ojima S, Shinjoh M, Tomiyama N;

DR WPI: 1998-195228/18.
DR N-PSDB: AAV29052.
XX
PT Recombinant Gluconobacter oxydans alcohol and/or aldehyde
PT dehydrogenase enzyme(s) - useful for converting substrate(s), e.g.
PT L-sorbose or D-sorbitol to 2-keto-L-gulonic acid
XX
PS Claim 1: Pages 38-40; 59pp; English.
XX
CC This is the amino acid sequence for the Gluconobacter oxydans
CC alcohol and/or aldehyde dehydrogenase A' enzyme. The enzymes
CC or recombinant organisms can be used to convert suitable substrates
CC to aldehydes, ketones or carboxylic acids, especially to convert
CC L-sorbose or D-sorbitol to 2-keto-L-gulonic acid, which can be
CC converted to L-ascorbic acid by standard procedures. The derivatives
CC of ADH enzymes have desired substrate specificity, higher affinity
CC to a substrate, lower affinity to an inhibitory compound, higher
CC stability against temperature and/or pH and higher catalytic speed.
XX
SQ Sequence 579 AA:

Query Match 90.1%; Score 640; DB 19; Length 579;
Best Local Similarity 91.9%; Pred. No. 6,1e-62;
Matches 124; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 1 MKPTSLMASAGALIALAPAFAYPTVDLLANPAGENISYSGONENYRHSPLTQIT 60
DB 1 mkptslmasagalaialapafayptvdellanpagsewisysqnenyrrhsplqit 60
QY 61 TENVGQLQVWARGMKGQKQVPTPLIHGVMYLANPGDVYQAIDAKTGDLIWEHRRQLPN 120
DB 61 tenvgqlqvwargmkgkqvptplihgvmylanpgdvyaiaidaktgdliwehrrqlpn 120
QY 121 IATINSFGPEPTRGMA 135
DB 121 iatinsfgeptrgma 135

RESULT 3

AAW37875
ID AAW37875 standard; Protein: 578 AA.

XX AAW37875;

DT 10-AUG-1998 (first entry)

DE Alcohol and/or aldehyde dehydrogenase A'' amino acid sequence.

XX Alcohol/aldehyde dehydrogenase A'' enzyme; recombinant organism;

KW aldehyde; ketone; carboxylic acid; L-sorbose; D-sorbitol;

XX 2-keto-L-gulonic acid; L-ascorbic; inhibition.

OS Gluconobacter oxydans.

XX Key Location/Qualifiers

FT Peptide 1..23

FT Protein /note= "signal peptide" 24..578

FT /note= "mature protein"

XX EP832974-A2.

XX 01-APR-1998.

XX 11-SEP-1997; 97EP-0115801.

XX 19-SEP-1996; 96EP-0115001.

XX (HOFF) HOFFMANN LA ROCHE & CO AG F.

XX Asakura A, Hoshino T, Ojima S, Shinjoh M, Tomiyama N;

DR WPI: 1998-195228/18.
 DR N-PSDB: AAV29053.
 XX
 PT Recombinant Gluconobacter oxydans alcohol and/or aldehyde
 PT dehydrogenase enzyme(s) - useful for converting substrate(s), e.g.
 PT L-sorbose or D-sorbitol to 2-keto-L-gulonic acid
 XX
 PS Claim 1; Pages 41-43; 59pp; English.
 XX
 CC This is the amino acid sequence for the Gluconobacter oxydans
 CC alcohol and/or aldehyde dehydrogenase A' enzyme. The enzymes
 CC or recombinant organisms can be used to convert suitable substrates
 CC to aldehydes, ketones or carboxylic acids, especially to convert
 CC L-sorbose or D-sorbitol to 2-keto-L-gulonic acid, which can be
 CC converted to L-ascorbic acid by standard procedures. The derivatives
 CC of ADH enzymes have desired substrate specificity, higher affinity
 CC to a substrate, lower affinity to an inhibitory compound, higher
 CC stability against temperature and/or pH and higher catalytic speed.
 CC
 SQ Sequence 578 AA;

Query Match 82.7%; Score 587.5; DB 19; Length 578;
 Best Local Similarity 83.0%; Pred. NO. 3.8e-56;
 Matches 112; Conservative 9; Mismatches 13; Indels 1; Gaps 1;
 QY 1 MKPTSLMASAGALALLAPAFQVTVTDELLANPAGEMISYCONQENYRHSPLTOIT 60
 Db 1 mklttllqssaaallvltipalag-taitcdemlanpagewinyngngenyhsplqt 59
 QY 61 TENNGQLQVLRMGQPKGVQVTPPLIHGVMYLANPGDVIOAIDAKTGDLWEHRRQLPN 120
 Db 60 adnvgqlqlvwarameagvqvtplhdgvmylanpgdvigaldqgdlwehrrqlpn 119
 QY 121 IATLNSFGEPTRGMA 135
 Db 120 lalinsfgeptrgma 134

RESULT 4
 AAW37876
 ID AAW37876 standard; Protein; 579 AA.
 XX
 AC AAW37876;
 XX
 DT 10-AUG-1998 (first entry)
 XX
 DE Alcohol and/or aldehyde dehydrogenase B amino acid sequence.
 XX
 KW Alcohol/aldehyde dehydrogenase B enzyme; recombinant organism;
 KW aldehyde; ketone; carboxylic acid; L-sorbose; D-sorbitol;
 KW 2-keto-L-gulonic acid; L-ascorbic; inhibition.
 XX
 OS Gluconobacter oxydans.
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..23
 FT /note="signal peptide"
 FT Protein 24..579
 FT /note="mature protein"
 XX
 PN EP832974-A2.
 XX
 PD 01-APR-1998.
 XX
 PF 11-SEP-1997; 97EP-0115801.
 XX
 PR 19-SEP-1996; 96EP-0115001.
 XX
 PA (HOFF) HOFFMANN LA ROCHE & CO AG F.
 XX
 PI Asakura A, Hoshino T, Ojima S, Shinjoh M, Tomiyama N;
 XX
 PT

DR WPI: 1998-195228/18.
 DR N-PSDB: AAV29054.
 XX
 PT Recombinant Gluconobacter oxydans alcohol and/or aldehyde
 PT dehydrogenase enzyme(s) - useful for converting substrate(s), e.g.
 PT L-sorbose or D-sorbitol to 2-keto-L-gulonic acid
 XX
 PS Claim 1; Pages 44-46; 59pp; English.
 XX
 CC This is the amino acid sequence for the Gluconobacter oxydans
 CC alcohol and/or aldehyde dehydrogenase B enzyme. The enzymes
 CC or recombinant organisms can be used to convert suitable substrates
 CC to aldehydes, ketones or carboxylic acids, especially to convert
 CC L-sorbose or D-sorbitol to 2-keto-L-gulonic acid, which can be
 CC converted to L-ascorbic acid by standard procedures. The derivatives
 CC of ADH enzymes have desired substrate specificity, higher affinity
 CC to a substrate, lower affinity to an inhibitory compound, higher
 CC stability against temperature and/or pH and higher catalytic speed.
 CC
 SQ Sequence 579 AA;

Query Match 79.9%; Score 567; DB 19; Length 579;
 Best Local Similarity 78.5%; Pred. NO. 7e-54;
 Matches 106; Conservative 13; Mismatches 16; Indels 0; Gaps 0;
 QY 1 MKPTSLMASAGALALLAPAFQVTVTDELLANPAGEMISYCONQENYRHSPLTOIT 60
 Db 1 mptlllrtsaavllltapaafqvtptidellanpagewinyngngenyhsplqt 60
 QY 61 TENNGQLQVLRMGQPKGVQVTPPLIHGVMYLANPGDVIOAIDAKTGDLWEHRRQLPN 120
 Db 61 adnvgqlqlvwarameagvqvtplhdgvmylanpgdvigaldqgdlwehrrqlpa 120
 QY 121 IATLNSFGEPTRGMA 135
 Db 121 valinaqgdtkryva 135

RESULT 5
 AAB35987
 ID AAB35987 standard; Protein; 754 AA.
 XX
 AC AAB35987;
 XX
 DT 01-MAR-2001 (first entry)
 XX
 DE Sorbitol dehydrogenase subunit 1 amino acid sequence.
 XX
 KW Sorbitol dehydrogenase; SDH; Gluconobacter oxydans; D-sorbitol;
 KW L-sorbose production; 2-keto-L-gulonic acid.
 XX
 OS Gluconobacter oxydans.
 XX
 PN WO200065066-A1.
 XX
 PD 02-NOV-2000.
 XX
 PF 23-APR-1999; 99WO-1B00736.
 XX
 PR 23-APR-1999; 99WO-1B00736.
 XX
 PA (CHOI/) CHOI E.
 PA (RHEE/) RHEE S.
 PA (LEE/) LEE E.
 XX
 PI Choi E, Rhee S, Lee E;
 XX
 DR WPI: 2000-687351/67.
 DR N-PSDB: AAC83153.
 XX
 PT Novel membrane-bound sorbitol dehydrogenase nucleic acid molecule
 PT isolated from Gluconobacter suboxydans useful for the fermentative

Query Match

22.3%; Score 158.5; DB 22; Length 443;

Best Local Similarity 32.0%; Pred. No. 6.4e-09;

Matches 47; Conservative 22; Mismatches 49; Indels 29; Gaps 7;

OY 6 LTMAS-----AGALALIAAPFAQVTPVDELNPAGEMISYQONENYRHSPLTQ 58
 DB 168 ltwagfndpqlngfslsadaipaea-ispvadq-----dwpaygrngsgqrfspikg 218
 OY 59 ITTENVGOLQLVMA-----RGKOPKRY--QVTPPLIHGVMYLANPGDVIOAIDAKTGD 109
 DB 219 lhadvnhlkeawvrtgdkpndgplnevpikvgdtlylctahgrlfaaldaasgk 278
 OY 110 LIWEHRQLPNATLNSFGEPT-RGMA 135
 DB 279 ekmyd---pelktnesifqhvrcryvs 302

RESULT 10

AAM95019
ID AAM95019 standard; Protein: 740 AA.

XX AC AAM95019;

DT 21-MAY-1999 (first entry)

DE Sorbitol dehydrogenase (SLDH) protein sequence.

XX Sorbitol dehydrogenase; SLDH; open reading frame; ORF2; L-sorbose;

KM D-sorbitol; vitamin C; enzyme.

OS Gluconobacter suboxydans.

FH Key Location/Qualifiers

FT Peptide 1..24

FT /note= "signal sequence"

FT Protein 25..740

FT /note= "mature protein"

XX EP897984-A2.

PD 24-FEB-1999.

PF 13-AUG-1998; 98EP-0115231.

PR 21-AUG-1997; 97EP-0114432.

XX (HOEF) HOFEMANN LA ROCHE & CO AG F.

PI Hosino T, Miyazaki T, Ojima S, Shinjoh M, Tomiyama N;

DR WPI; 1999-134646/12.

DR N-PSDB; AAX21501.

PT New D-sorbitol dehydrogenase gene and recombinant protein - useful
for production of L-sorbose, an intermediate in vitamin C production

PS Claim 1; Fig 3A-D; 39pp; English.

XX This represents a sorbitol dehydrogenase (SLDH) protein. The DNA
 CC encoding the SLDH enzyme also encodes an open reading frame (ORF2)
 CC product upstream of the SLDH open reading frame, needed for SLDH activity
 CC in vivo. Host cells transformed by a vector comprising the SLDH DNA
 CC sequence are used for the recombinant expression of the sorbitol
 CC dehydrogenase. Recombinant D-sorbitol dehydrogenase is useful for
 CC producing L-sorbose from D-sorbitol. L-sorbose is an important
 CC intermediate in vitamin C production.

SO Sequence 740 AA;

Query Match 20.5%; Score 145.5; DB 20; Length 740;

Best Local Similarity 34.0%; Pred. No. 3.5e-07;

Matches 34; Conservative 25; Mismatches 30; Indels 11; Gaps 5;

OY 23 AQTPTVDELLANPAGEMISYQONENYRHSPLTQITTEENVGOLQ--VMARGM--QNG 78
 DB 86 sqypamapqgsanpargdwvaygrddhqrtyrslsetlpensasklvatvhtgysyprg 145
 OY 79 KV-----QVTPPL-IHOGVMYLANPGDVIOAIDAKTGDLLW 112
 DB 146 qvkwaaettpikvgdglyrcsamndlik-10patgkqlw 184

RESULT 11

ABG25909
ID ABG25909 standard; Protein: 948 AA.

XX AC ABG25909;

DT 18-FEB-2002 (first entry)

DE Novel human diagnostic protein #25900.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KM food supplement; medical imaging; diagnostic; genetic disorder.

OS Homo sapiens.

PN W0200175067-A2.

PD 11-OCT-2001.

PF 30-MAR-2001; 2001WO-US08631.

PR 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

XX (HYSE-) HYSEQ INC.

PI Drmanac RT, Liu C, Tang YT;

DR WPI; 2001-639362/73.

DR N-PSDB; AAS90096.

PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity

XX Claim 20; SEQ ID No 56268; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.

CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

SO Sequence 948 AA;

PF treat Haemophilus influenzae infection or associated disease, e.g.,
 PF otitis media and conjunctivitis -
 XX
 XX
 PS Disclosure: Fig 1: 29pp: English.
 CC This sequence represents the major outer membrane protein P5 of
 CC non-typeable H. influenzae. One or more surface exposed loops of this
 CC protein may be replaced with a modified peptide of the invention. Each
 CC of these peptides contain an LRI(f) peptide which is a 19 amino acid
 CC peptide derived from the sequence of MOMP P5 from strain nH1128,
 CC representing amino acids Arg117 to Gly135. This peptide represents the
 CC third exposed loop of P5 and is a potential B cell epitope. The loops
 CC of the invention are modified in terms of being in a non-native
 CC environment in the recombinant outer membrane protein. The modified
 CC MOMP P5 may be used to induce an immune response in a mammal to
 CC prevent or treat Haemophilus influenzae infection or associated
 CC disease, e.g., otitis media, sinusitis, conjunctivitis, or lower
 CC respiratory tract infection.
 CC
 SQ Sequence 353 AA;

Query Match 10.9%; Score 77.5; DB 22; Length 353;
 Best Local Similarity 24.4%; Pred. No. 4.1;
 Matches 39; Conservative 21; Mismatches 71; Indels 29; Gaps 5;

OY 1 MKPSTLWASGALALAPAFQVTPYTDLELANPAG-----WISYGO 46
 DB 1 mktstlwlavag-----laasvagaapqentfyagvkagafndglralareykvgyhr 56
 OY 47 NOENRHSPLQITTE-----VGQLQLVWARGMPGKQVTPPLHHDGVMYLANPGDV 99
 DB 57 nsftgyvfgygqllngnmlglavelyddfgkrakgrevkvtvthhthgltslksyev 116
 OY 100 IQAIPD-AKTGDLI--WEHRQLEPNIAIATLNSFGEPTGMA 135
 DB 117 legldvygkagvalrvsdykllynnensclkklghehtrara 156

RESULT 14

AAW09614
 ID AAW09614 standard; Protein: 824 AA.

AAW09614;

20-MAY-1997 (first entry)

Pathogenic Staphylococcus aureus ORF1 gene product.

Pathogenicity; vaccine; immunoassay; antibody.

Staphylococcus aureus strain NCTC 8325-4 (ISP8).

W09641878-A1.

27-DEC-1996.

13-JUN-1996; 96WO-AU00353.

13-JUN-1995; 95AU-0003507.

(AUSU) UNITV AUSTRALIAN NAT.

Adham El-Adhami W, Matthei KI, Stewart PR;

WPI: 1997-065465/06.

N-PSDB; AAT47757.

Nucleic acids specific to pathogenic Staphylococcus aureus strains -

for diagnosis and immunogenic protein prodn.

Claim 30, Page 62-65; 88pp: English.

CC An 824-amino acid protein (AAW09614) and 192-amino acid protein
 CC (AAW09615) were identified as products of 2 open reading frames in
 CC a nucleic acid (AAT47757) isolated from pathogenic Staphylococcus
 CC aureus strain ISP8. The 2 gene products are highly hydrophilic
 CC and have high surface probability and antigenic index measurements,
 CC suggesting that they are surface expressed. The polypeptides,
 CC their fragments, and antibodies raised against them can be used in
 CC vaccines and as immunoassay reagents for detecting pathogenic S.
 CC aureus strains.
 CC
 SQ Sequence 824 AA;

Query Match 10.7%; Score 76; DB 18; Length 824;
 Best Local Similarity 24.8%; Pred. No. 19;
 Matches 33; Conservative 18; Mismatches 34; Indels 48; Gaps 7;

OY 19 APAPFAQVT-----PYTDLELANPAGEWISYGOENRHSPLQITTEVGOQLIV 70
 DB 683 apgtkvtregqgkxklttptlknplltgelskgskeeltkdpineit----- 732
 OY 71 WARGMPGKQVTPPLHD-----GVWYLANP--GDVIO-AIDAKT----- 107
 DB 733 -----eygpetltpnrhdefdpxlpcgkeevygpjknpcbtgvdvrvpvdsvtkygpv 787
 OY 108 -GDVMEHRROLP 119
 DB 788 kgdsive-keelp 799

RESULT 15

ABB59208
 ID ABB59208 standard; Protein: 1419 AA.

ABB59208;

26-MAR-2002 (first entry)

Drosophila melanogaster polypeptide SEQ ID NO 4416.

Drosophila; developmental biology; cell signalling; insecticide;

pharmaceutical.

Drosophila melanogaster.

W0200171042-A2.

27-SEP-2001.

23-MAR-2001; 2001WO-US09231.

23-MAR-2000; 2000US-191637P.

11-JUL-2000; 2000US-0614150.

(PEKE) PE CORP NY.

Venter JC, Adams M, Li PMD, Myers EW;

WPI: 2001-656860/75.

N-PSDB; ABL03311.

New isolated nucleic acid detection reagent for detecting 1000 or more

genes from Drosophila and for elucidating cell signalling and cell-cell

interactions -

Disclosure: SEQ ID NO 4416; 21pp + Sequence Listing: English.

The invention relates to an isolated nucleic acid detection reagent

capable of detecting 1000 or more genes from Drosophila. The invention is

useful in developmental biology and in elucidating cell signalling and

cell-cell interactions in higher eukaryotes for the development of

insecticides, therapeutics and pharmaceutical drugs. The invention

discloses genomic DNA sequences (ABLI6176-ABLI30511), expressed DNA

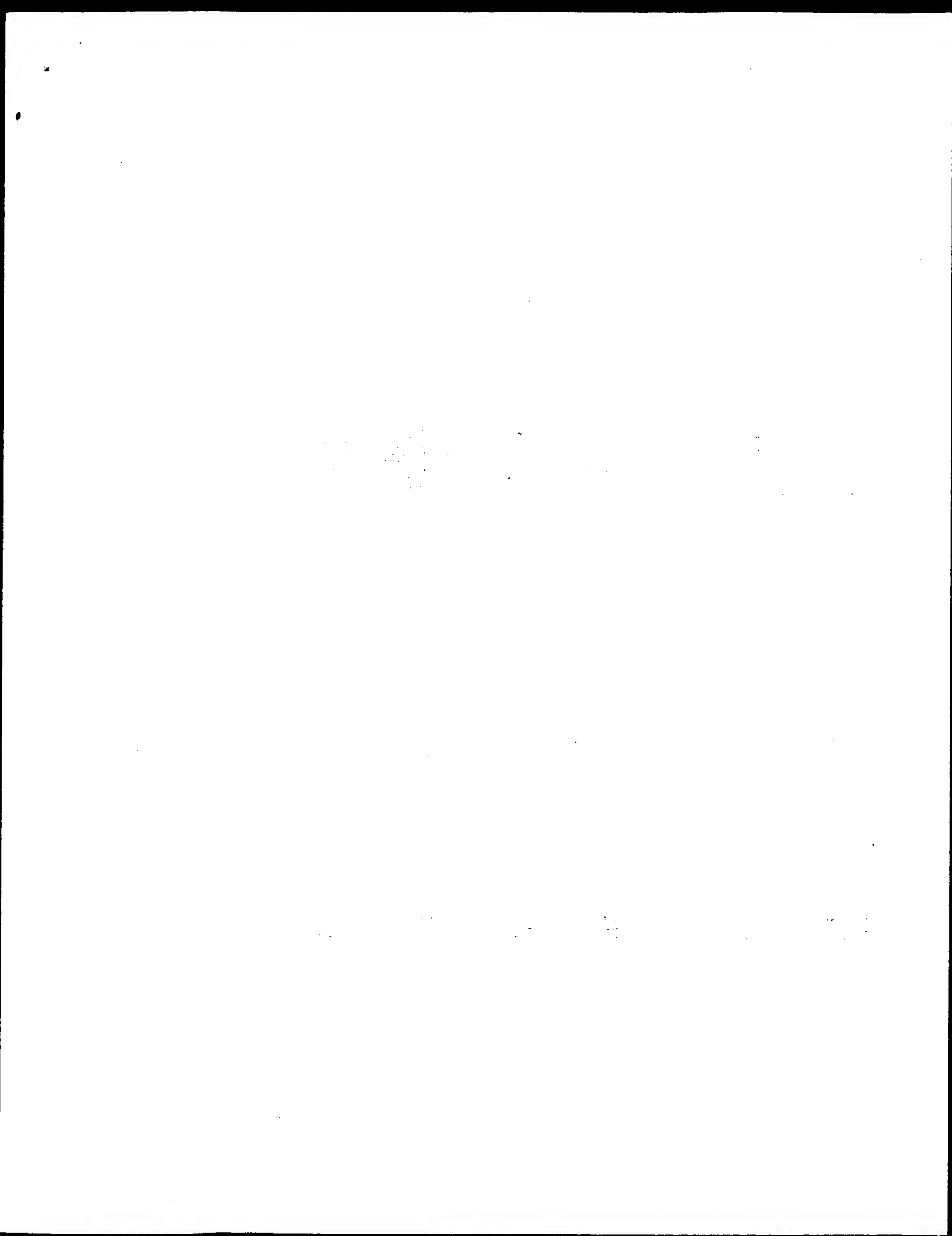
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (AB057737-AB072072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX

SO Sequence 1419 AA;

Query Match 10.6%; Score 75.5; DB 22; Length 1419;
Best Local Similarity 22.6%; Pred. No. 46;
Matches 30; Conservative 15; Mismatches 65; Indels 23; Gaps 3;

QY 9 ASAGALALLAPAFQVTPVTDELIANPAGEWISYGN-----QENYRHSPLTQITTE 62
Db 716 asrtatlaaaaaaagltptatvaaaatppaasakafndvssfsstnkrhgatssnmr 775
QY 63 NVGQLQLVWARGMQPKVQVTPLIHDGVWYLANPGDVIQAIDAKTGDLIWEHRRQLPNTA 122
Db 776 ssnqaaaaaaatkaaalatvp-----aapngllscrksns-----qnsplya 818
QY 123 TLNSGEPTRGMA 135
Db 819 tpxltltptpfia 831

Search completed: May 24, 2002, 10:05:02
Job time: 241 sec



GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2002, 10:02:57 ; Search time 44.35 Seconds

(without alignments)
74.351 Million cell updates/sec

Title: US-08-934-506a-5_COPY_1_135

Perfect score: 1 MKPISLMAAGALALAAP.....ROLPNITLNSFGPTKMA 135

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 2442594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5a_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5b_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6a_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6b_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS.COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	193.5	27.3	754	4	US-09-296-284-4
2	187	26.3	738	1	US-07-985-458-3
3	184.5	26.0	720	4	US-09-296-284-25
4	145.5	20.5	739	3	US-09-136-251-2
5	72.5	10.2	462	4	US-08-630-915A-38
6	70.5	9.9	540	1	US-08-368-071-11
7	70.5	9.9	540	1	US-08-458-181-11
8	70.5	9.9	540	5	PCT-US93-02172-11
9	68	9.5	159	3	US-08-796-792-2
10	67.5	9.5	851	1	US-08-369-796-2
11	67.5	9.5	851	2	US-08-852-091-2
12	67.5	9.5	851	2	US-08-820-754-2
13	67.5	9.5	851	3	US-08-956-652-2
14	67.5	9.5	851	3	US-08-956-869-2
15	67.5	9.5	851	3	US-09-012-710-2
16	67.5	9.5	851	3	US-08-948-547-2
17	67.5	9.5	851	4	US-09-087-465-4
18	67.5	9.5	851	4	US-09-364-970-2
19	67.5	9.5	851	4	US-09-556-273-2
20	67.5	9.5	851	4	US-08-956-653A-2
21	67.5	9.5	851	5	PCT-US95-17025-2
22	67.5	9.5	852	1	US-08-276-099A-13
23	67.5	9.5	852	1	US-08-781-890-13
24	67.5	9.5	974	2	US-08-868-786-6
25	66.5	9.4	210	4	US-09-247-155-117
26	65.5	9.2	876	2	US-08-510-215A-2
27	65.5	9.2	1229	3	US-09-310-293-2

28	65.5	9.2	1229	4	US-09-579-376-2	Sequence 2, Appli
29	64.5	9.1	286	4	US-08-964-127-4	Sequence 4, Appli
30	64.5	9.1	286	4	US-09-496-692-4	Sequence 1, Appli
31	64.5	9.1	520	4	US-08-964-127-2	Sequence 2, Appli
32	64.5	9.1	520	4	US-09-496-692-2	Sequence 2, Appli
33	64.5	9.1	523	2	US-08-923-772-2	Sequence 2, Appli
34	64.5	9.1	523	2	US-09-385-287-2	Sequence 2, Appli
35	64	9.0	966	2	US-08-868-786-2	Sequence 2, Appli
36	63.5	8.9	777	2	US-08-874-678-3	Sequence 3, Appli
37	63.5	8.9	777	3	US-08-643-839-3	Sequence 3, Appli
38	63.5	8.9	820	4	US-09-313-677-21	Sequence 21, Appli
39	63.5	8.9	926	4	US-09-313-677-2	Sequence 2, Appli
40	63.5	8.9	933	4	US-09-313-677-19	Sequence 19, Appli
41	63.5	8.9	967	4	US-09-313-677-17	Sequence 17, Appli
42	63.5	8.9	1298	1	US-08-222-616-33	Sequence 33, Appli
43	63.5	8.9	1298	1	US-08-340-011-2	Sequence 2, Appli
44	63.5	8.9	1298	3	US-08-901-710-2	Sequence 2, Appli
45	63.5	8.9	1298	5	PCT-US95-04228-33	Sequence 33, Appli

ALIGNMENTS

```
RESULT 1
US-09-296-284-4
; Sequence 4, Application US/09296284A
; Patent No. 6204040
; GENERAL INFORMATION:
; APPLICANT: Choi, Eun-Sung
; APPLICANT: Rhee, Sang-Ki
; APPLICANT: Lee, Eun-Hae
; TITLE OF INVENTION: Glucanobacter Suboxydans Sorbitol Dehydrogenase, Genes
; TITLE OF INVENTION: and Methods of Use Thereof
; FILE REFERENCE: 1533.0870000
; CURRENT APPLICATION NUMBER: US/09/296,284A
; CURRENT FILING DATE: 1999-04-22
; NUMBER OF SEQ ID NOS: 87
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 754
; TYPE: PRT
; ORGANISM: Glucanobacter suboxydans
US-09-296-284-4

Query Match      27.3%; Score 193.5; DB 4; Length 754;
Best Local Similarity 37.4%; Pred. No. 6.4e-14;
Matches 49; Conservative 17; Mismatches 56; Indels 9; Gaps 3;

QY 14 LALLAAPAFQVTPV-----TDELLANPPAGEWISTGQONENTRHSPLDITTEWNGQ 66
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 17 LCGAALACATSPVALADDTGATITNADQHPGDMMSYCRYSQRYSPDITRDMSN 76

QY 67 LQLYMARQGPQVQVQV-TPLIHDGVWYLANPGDVIQADIKGLIWEHRDLP-NIATL 124
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 77 LKIAWYDDITDNGOGGTPLIVGVWATTTNMSKMAIDAAIGKLWSDPKVPENIADR 136

QY 125 NSRGEPTRGMA 135
DB 137 GCCDYNRGAA 147

RESULT 2
US-07-985-458-3
; Sequence 3, Application US/07985458
; Patent No. 5344777
; GENERAL INFORMATION:
; APPLICANT: Tamaki, Toshimi;
; APPLICANT: Takemura, Hiroshi;
; APPLICANT: Takemura, Kenji;
; APPLICANT: Fukaya, Masahiro;
; APPLICANT: Okumura, Hajime and
; APPLICANT: Kawamura, Yoshiya
```

Query Match	26.3%;	Score 187;	DB 1;	Length 738;
Best Local Similarity	32.4%;	Pred. No. 3.5e-13;		
Matches 44;	Conservative 23;	Mismatches 59;	Indels 10;	Gaps 3

Query Match	26.0%;	Score 184.5;	DB 4;	Length 720;
Best Local Similarity	38.4%;	Pred. No. 6.6e-13;		
Matches 43;	Conservative 18;	Mismatches 46;	Indels 5;	Gaps 3

RESULT 4
 US-09-136-251-2
 : Sequence 2, Application US/09136251A
 : Patent No. 6127156
 :
 : GENERAL INFORMATION:
 : APPLICANT: HOSHINO, Tatsuo
 : APPLICANT: MIYAZAKI, Taro
 : APPLICANT: OJIMA, Setsuko
 : APPLICANT: SHINJOH, Masako
 : APPLICANT: TOMIYAMA, No. 6127156ibumi
 : TITLE OF INVENTION: D-SORBITOL DEHYDROGENASE GENE
 : FILE REFERENCE: D-Sorbitol Dehydrogenase Gene
 : CURRENT APPLICATION NUMBER: US/09/136,251A
 : CURRENT FILING DATE: 1998-08-19
 : EARLIER APPLICATION NUMBER: EP 97114432.4
 : EARLIER FILING DATE: 1997-08-21
 : NUMBER OF SEQ ID NOS: 9
 : SOFTWARE: PatentIn Ver. 2.1
 : SEQ ID NO 2

```

; ORGANISM: Gluconobacter suboxydans
; FEATURE:
; NAME/KEY: SIGNAL
;

```


STREET: 4370 LA JOLLA VILLAGE DRIVE, SUITE 700
CITY: SAN DIEGO
STATE: CALIFORNIA
COUNTRY: UNITED STATES
ZIP: 92122
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/458,181
FILING DATE: 30-DEC-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: CAMPBELL, CATHERIN
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-LJ 1275
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-535-9001
TELEFAX: 619-535-8949
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 540 amino acids
Type: amino acid
TOPOLOGY: linear
US-08-458-181-11

Query Match	9.9%	Score 70.5;	DB 1;	Length 540;
Best Local Similarity	21.5%;	Pred. No. 6.5;		
Matches 31; Conservative	22;	Mismatches 48;	Indels 43;	Gaps 4;

[illegible]

RESULT 8
 PCT-US93-02172-11
 Sequence 11, Application PC/TUS9302172
 GENERAL INFORMATION:
 APPLICANT: La Jolla Cancer Research Foundation
 TITLE OF INVENTION: RECOMBINANT CALF INTESTINAL ALKALINE
 PHOSPHATASE
 NUMBER OF SEQUENCES: 13
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: La Jolla Cancer Research Foundation
 STREET: 10901 North Torrey Pines Road
 CITY: La Jolla
 STATE: California
 COUNTRY: USA
 ZIP: 92037
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version 1.25 (EPO)
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: PCT/US93/02172
 FILING DATE: 19930310
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/07/849,219
 FILING DATE: 10-MAR-1992

```

1 TELECOMMUNICATION INFORMATION
2 TELEPHONE: (619) 455-6480
3 TELEFAX: (619) 455-0181
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1
```

Query Match	9.9%;	Score 70.5;	DB 5;	Length 540;
Best Local Similarity	21.5%;	Pred. No. 6.5;		
Matches	31;	Conservative	22;	Mismatches 48;
				Indels 43;
				Gaps 4;

```

QY 18 AARAFQAVPPVYDELLANPPEGEMISTVGQNDENRHH-----SPLTQ 58
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 304 ADSLSEMEVEAVRLLSRNPQGFYLLFVEGGRIQDGHGAGTALATLEAVMEDSALEKASQ 365
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 59 ITTENGGQTLQVWA-----RG-----MPSKVQVPLIHIDGVMYLAN 95
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 364 LTNEK-DITLLTIAHSHVFAFGGTYTGTSTIFGLAPLNKQDSKSTSLIYNGPGVYLN 422
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 96 PGDIVQAIIDAKTGDTLWEHRRLP 119
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 423 SGRRPWTDAESGDVNYKQQAAYP 446
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

RESULT 9
US-08-796-792-2
; Sequence 2, Application US/08796792
Datafile US-08796792-2

GENERAL INFORMATION:
APPLICANT: Gennaro, Maria L.
APPLICANT: Iyashchenko, Konstantin P.
APPLICANT: Manca, Claudia M.A.
TITLE OF INVENTION: MYCOBACTERIUM TUBERCULOSIS
TITLE OF INVENTION: SPECIFIC PROTEINS AND GENES, MIXTURES OF ANTIGENS AND USES
TITLE OF INVENTION: THEREOF
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson, P.C.
STREET: 45 Rockefeller Plaza, Suite 2800
CITY: New York
STATE: NY
COUNTRY: US
ZIP: 10111
COMPUTER READABLE FORM:
MEDIUM TYPE: diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows95
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/796,792
FILING DATE: 06-FEB-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/011,364
FILING DATE: 09-FEB-1996
ATTORNEY/AGENT INFORMATION:
NAME: Hone, William J.
REGISTRATION NUMBER: 26,739
REFERENCE/DOCKET NUMBER: 07763/03301
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-765-5070
TELEFAX: 212-258-2291
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 159 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein

Fri May 24 11:27:18 2002

us-08-934-506a-5_copy_1_135.rai

Page 8

Db 568 LKDLMDGRIMGVSRS-----QERRLLKKTMSGTFLLRFSSESSEG 608

Search completed: May 24, 2002, 10:02:58
Job time: 232 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2002, 10:06:04 ; Search time 56.04 Seconds
(without alignments)
231.479 Million cell updates/sec

Title: US-08-934-506a-5_COPY_1_135

Perfect score: 710
Sequence: 1 MKPTSLMASAGALALAAP.....RQLPNIALNSGEPTKMA 135

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: PIR1:
2: PIR2:
3: PIR3:
4: PIR4:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	191	26.9	742	2 A49340	alcohol dehydrogen
2	187	26.3	738	2 S14270	alcohol dehydrogen
3	179	25.2	742	2 JS0326	alcohol dehydrogen
4	165.5	22.3	708	2 S52317	quinohemoprotein e
5	158.5	22.3	796	1 JV0107	glucose dehydrogen
6	158.5	22.3	796	2 H90644	glucose dehydrogen
7	158.5	22.3	796	2 H85495	glucose dehydrogen
8	157.5	22.2	796	2 AG0523	glucose dehydrogen
9	148	20.8	778	2 G98221	glucose dehydrogen
10	148	20.8	778	2 A13064	glucose dehydrogen
11	143	20.1	803	2 F83360	glucose dehydrogen
12	140.5	19.8	601	2 E95863	alcohol dehydrogen
13	137	19.3	623	2 B83399	quinoprotein alcoh
14	134	18.9	808	1 OPRKX	glucose dehydrogen
15	123	17.3	801	1 S00943	glucose dehydrogen
16	107	15.1	809	2 B98314	probable quinolate d
17	107	15.1	809	2 A12968	hypothetical prote
18	101.5	14.3	626	2 JU0706	alcohol dehydrogen
19	99	13.9	809	2 A55547	quinolate-shikimate
20	93.5	12.1	639	2 JC4810	polyvinyl-alcohol
21	86	11.8	613	2 F69424	conserved hypotet
22	83.5	11.8	2802	2 F97686	cyclic beta-(1-2)
23	83.5	11.8	2831	2 A12911	beta (1-->2) gluc
24	82	11.5	666	2 C75264	probable serine/th
25	81.5	11.5	393	2 AD0350	probable lipoprote
26	80.5	11.3	647	2 C83012	probable chemotaxi
27	80.5	11.3	772	2 H86016	hypothetical prote
28	78.5	11.1	148	1 S05638	cuticle protein 8
29	78.5	11.1	509	2 C70624	probable two compo

ALIGNMENTS

30	78	11.0	207	2 AH1548	3-methyladenine DN
31	77.5	10.9	353	2 C64187	outer membrane pro
32	77	10.8	733	2 S78376	photosystem I p700
33	76.5	10.8	224	2 AD2115	phosphoribosylform
34	76.5	10.8	400	2 AD0604	D-alanyl-D-alanine
35	76	10.7	306	2 D75562	N-acetyl-gamma-glu
36	76	10.7	386	2 A82284	conserved hypotet
37	76	10.7	573	2 S68591	methanol dehydroge
38	76	10.7	580	2 A80994	gamma-glutamyltran
39	75.5	10.6	407	2 C82804	conserved hypotet
40	75.5	10.6	684	2 T36771	probable integral
41	75.5	10.6	3972	2 S75251	hypothetical prote
42	74.5	10.5	969	2 A75634	MCRB-related prote
43	74	10.4	380	2 C83171	conserved hypotet
44	73.5	10.4	347	2 A83617	probable oxidoredu
45	73.5	10.4	584	2 AH2506	ser/thr protein Ki

RESULT 1
A49340
alcohol dehydrogenase (EC 1.1.1.-) precursor - Acetobacter pasteurianus (strain NC113
C:Species: Acetobacter pasteurianus
C:Date: 07-Apr-1994 #sequence-revision 07-Apr-1994 #text-change 20-Jun-2000
C:Accession: A49340
R: Takemura, H.; Kondo, K.; Horinouchi, S.; Beppu, T.
J. Bacteriol. 175, 6857-6866, 1993
A:Title: Induction by ethanol of alcohol dehydrogenase activity in Acetobacter pasteu
A:Reference number: A49340; MUID:94042848
A:Accession: A49340
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-742 <TAK>
A:Cross-references: GB:DJ1893; NID:9517067; PIDN:BAA40252.1; PID:9452586
C:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)
C:Keywords: alcohol metabolism; NAD; oxidoreductase

Query Match 26.9%; Score 191; DB 2; Length 742;
Best Local Similarity 37.0%; Pred. No. 5.8e-11;
Matches 47; Conservative 18; Mismatches 58; Indels 4; Gaps 3;
QY 11 AGALLAAPAFQVPTVDL--ANPPAGEMISGQNGENTRHSPLOTITENVGOLQ 68
DB 23 AALPYAAVPARADCGGNTGEALIHADDPENWLSYGRYSERQYSPLOINRNVGDLK 82
QY 69 LVNARGMQEGKV-VTPLIHDGVMYLANPGDVTQATDAGTGLIWEHRQLP-NIATLNS 126
DB 83 LAWYTLDTNRGQCATPLVVDGIMVATWMSKMEALDAIGKILWOYDPKVGNIADKGC 142
QY 127 FGEPTRG 133
DB 143 CDTVNRG 149

RESULT 2
S14270
alcohol dehydrogenase, membrane-bound (EC 1.1.1.-) 72K chain precursor - Acetobacter
C:Species: Acetobacter polyoxogenes
A:Variety: Strain NB11028
C:Date: 21-Nov-1993 #sequence-revision 26-Jul-1996 #text-change 20-Jun-2000
C:Accession: S14270
R: Tamaki, T.; Fukaya, M.; Takemura, H.; Tayama, K.; Okumura, H.; Kawamura, Y.; Nishiy
Biochim. Biophys. Acta 1088, 292-300, 1991
A:Title: Cloning and sequencing of the gene cluster encoding two subunits of membrane
A:Reference number: S14270; MUID:91159482
A:Accession: S14270
A:Molecule type: DNA
A:Residues: 1-738 <TAM>
A:Cross-references: GB:D00635; NID:9216185; PIDN:BAA00528.1; PID:9216186
A:Experimental source: strain NB11028

[illegible]

C:\Genetics:
A:\Gene: EGS0128
C:\Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)

```
Query Match          22.3%; Score 158.5; DB 2; Length 796;  
Best Local Similarity 32.0%; Pred. No. 1.1e-07;  
Matches 47; Conservative 22; Mismatches 49; Indels 29; Gaps 7;
```

OY 6 LHMAS-----AGALLALAAPAFQAQVPTVDDELIANPAGGEMISTYGONENYRHSPLTQ 58
| | | | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 135 LTWAFNDPOEINGTLISDADTPAEA-ISPVADQ-----DWPAYGRNQGQRSPSLKQ 185
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

OY 59 ITTENVGOLQLVMA-----RGMPQKV--QVPTPLIHGVMYLANPCDVIAQIDARTGD 109
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 186 INADVNHNLKEAWVFRTGDVKQPNDPGETITNEVTPIKVGDTLLYCTAHQRFLDALDAASGK 245
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

OY 110 LIWEHRROLPNATLNSFGEPT-RGMA 135
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 246 EKMHYD---PELKTNESFOHVCRCGVS 269

RESULT 7
H85495
glucose dehydrogenase [imported] - Escherichia coli (strain O157:H7, substrain EDL93333)
C:\Species: Escherichia coli
C:\Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
C:\Accession: H85495
J.Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; May
Miller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Diallanta, E.; Potamoulsis, K.; Apodaca
Nature 409, 529-533, 2001
A:\Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:\Reference number: A85480; MUID:21074935; PMID:11206551
A:\Accession: H85495
A:\Status: preliminary
A:\Molecule type: DNA
A:\Residues: 1796 <STO>
A:\Cross-references: GB:AEO05174; NID:g12512839; PIDN:AAG54428.1; GSFD:GN00145; UWGP:
A:\Experimental source: strain O157:H7, substrain EDL933
C:\Genetics:
A:\Gene: gcd
C:\Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)

Query Match 22.3%; Score 158.5; DB 2; Length 796;
Best Local Similarity 32.0%; Pred. No. 1.1e-07;
Matches 47; Conservative 22; Mismatches 49; Indels 29; Gaps 7;

OY 6 LHMAS-----AGALLALAPAFQAQVPTVDDELIANPAGGEMISTYGONENYRHSPLTQ 58
| | | | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 135 LTWAFNDPOEINGTLISDADTPAEA-ISPVADQ-----DWPAYGRNQGQRSPSLKQ 185
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

OY 59 ITTENVGOLQLVMA-----RGMPQKV--QVPTPLIHGVMYLANPCDVIAQIDARTGD 109
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 186 INADVNHNLKEAWVFRTGDVKQPNDPGETITNEVTPIKVGDTLLYCTAHQRFLDALDAASGK 245
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

OY 110 LIWEHRROLPNATLNSFGEPT-RGMA 135
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 246 EKMHYD---PELKTNESFOHVCRCGVS 269

RESULT 8
AG0523
glucose dehydrogenase [imported] - Salmonella enterica subsp. enterica serovar Typh1
C:\Species: Salmonella enterica subsp. enterica serovar Typh1
A:\Note: this species has also been called Salmonella typhi
C:\Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 27-Nov-2001
C:\Accession: AG0523
R.Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Main, J.; Church
Th, T.; Connerton, P.; Cronin, A.; Davies, P.; Davies, R.M.; Dowd, L.; White, N.; Farr
S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A:\Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens,

RESULT 15

500943

glucose dehydrogenase (pyrroloquinoline-quinone) (EC 1.1.99.17) - *Acinetobacter calcoace*
C:Species: *Acinetobacter calcoaceticus*

C; Date: 10-Sep-1999 #sequence_revision

```
C:\Accession: S00843 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
```

C; ACCESSION: 500943
R; Cleton-Jansen, A.

Nucleic Acids Res. 16, 6228, 1988

Nucleic Acids Res. 16, 6228, 1988

A;Title: Nucleotide sequence of the gene coding for quinoprotein glucose dehydrogenase from *Paracoccus denitrificans* strain 20013

A; Reference number: S00943; MUID: 88289368

A;Accession: 500943
A:Molecule type: DN

A;Residues: 1-801 <C

A;Cross-references: EMB

A; Experimental source: strain LMD 79.41

A: Description: catalyzes the conversion of

C; Function:

A:Pathway: respiratory chain

C; Superfamily: glucose dehydrogenase

C; Keywords: oxidoreductase; pyrroloquinoline quinone; respiratory

F:9-35/Domain: transmembrane #status predicted <TM1>

transmembrane	#status	predicted	<TM2>
F;39-3//Domain:	transmembrane	predicted	<TM2>
F:61-79/Domain:	transmembrane	predicted	<TM3>

F;94-108/Domain: transmembrane #status predicted <TM4

F:118-137/Domain: transmembrane #status predicted <TM5

F;91,93/Binding site: ublignone (Arg, Asp) #status pr

F:\4\1\Active site: Asp #status predicted

Query Match	Score
-------------	-------

Best Local

Matches 39; Conservative 16; Mismatches 48; Indels .

0V 20 PAFAOVTPYTDL LNPAGEWTSYGNONEVRHSPITOTT

Case	Age	Sex	Occupation	Duration of illness	Site of lesion	Pathological changes	Microscopic findings	Diagnosis
1	45	M	Farmer	10 years	Brain	Chronic	Microscopic findings	Chronic
2	55	F	Housewife	5 years	Brain	Chronic	Microscopic findings	Chronic
3	65	M	Teacher	15 years	Brain	Chronic	Microscopic findings	Chronic
4	75	F	Retired	20 years	Brain	Chronic	Microscopic findings	Chronic
5	85	M	Farmer	25 years	Brain	Chronic	Microscopic findings	Chronic
6	95	F	Housewife	30 years	Brain	Chronic	Microscopic findings	Chronic
7	105	M	Farmer	35 years	Brain	Chronic	Microscopic findings	Chronic
8	115	F	Housewife	40 years	Brain	Chronic	Microscopic findings	Chronic
9	125	M	Farmer	45 years	Brain	Chronic	Microscopic findings	Chronic
10	135	F	Housewife	50 years	Brain	Chronic	Microscopic findings	Chronic
11	145	M	Farmer	55 years	Brain	Chronic	Microscopic findings	Chronic
12	155	F	Housewife	60 years	Brain	Chronic	Microscopic findings	Chronic
13	165	M	Farmer	65 years	Brain	Chronic	Microscopic findings	Chronic
14	175	F	Housewife	70 years	Brain	Chronic	Microscopic findings	Chronic
15	185	M	Farmer	75 years	Brain	Chronic	Microscopic findings	Chronic
16	195	F	Housewife	80 years	Brain	Chronic	Microscopic findings	Chronic
17	205	M	Farmer	85 years	Brain	Chronic	Microscopic findings	Chronic
18	215	F	Housewife	90 years	Brain	Chronic	Microscopic findings	Chronic
19	225	M	Farmer	95 years	Brain	Chronic	Microscopic findings	Chronic
20	235	F	Housewife	100 years	Brain	Chronic	Microscopic findings	Chronic

Db 152 PETAQAVPGVAE-----SDWPAYGRTQAGVRYSP LKQINDQNVKDLKVAWT--LRTGD 202

[illegible]

xy 80 V-----QVIFLHDSVMYLANPGDV IQALDAKIGDLLWEHRRQLPNIALNSFC 128
|||: |*: ||||

Db 203 LKTDNDSGEITNOVPIKIGNMFEICTAHOOLAIDPATGKEKW---BEDPKIKTDKSFQ 259

Search completed: May 24, 2002, 10:06:05
Job time: 289 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2002, 10:15:49 ; Search time 27.86 Seconds
(without alignments)
187,622 Million cell updates/sec

Title: US-08-934-506a-5_COPY_1_135

Perfect score: 710
Sequence: 1 MKPTSLMASAGALALAAP.....ROLPNATLNSFGPRTGMA 135

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	211	29.7	757	1 DHET_GLUOX	005542 gluconobact
2	187	26.3	738	1 DHET_ACEBO	P28036 acetobacter
3	186	26.2	739	1 DHET_ACEBO	044002 acetobacter
4	179	25.2	742	1 DHET_ACEAC	P18278 acetobacter
5	158.5	22.3	796	1 EXAA_PSEAE	092477 pseudomonas
6	137	19.3	623	1 DHG_GLUOX	P21715 gluconobact
7	134	18.9	808	1 DHG_GLUOX	P05455 acinetobact
8	123	17.3	801	1 DHM1_PARDE	P12283 paracoccus
9	119.5	16.8	631	1 DHM1_PARDE	P15279 methyllobact
10	113.5	16.0	626	1 DHM1_METEX	P16027 methyllobact
11	101.5	14.3	626	1 XOXE_PARDE	P29968 paracoccus
12	99	13.9	809	1 GUN2_THIEFU	059086 acetobacter
13	99	13.9	809	1 GUN2_THIEFU	092478 xanthomonas
14	91.5	12.9	790	1 QUIA_XANCI	P11734 locusta mig
15	78.5	11.1	148	1 CU08_LOCMT	092089 listeria in
16	78	11.0	207	1 3MGH_LISIN	P43840 haemophilus
17	77.5	10.9	353	1 OM51_HABIN	P49480 odontella s
18	77	10.8	733	1 PSAB_ODOSI	035734 marmota mon
19	76	10.7	233	1 TNFA_MARMO	P38539 methyllobil
20	76	10.7	571	1 DHM1_METEX	092472 heterocapsa
21	73	10.3	776	1 PSAB_HETTR	P15683 rattus norv
22	70.5	9.9	540	1 PPBI_RAT	P58621 listeria mo
23	70	9.9	207	1 3MGH_LISMO	P26222 thermomonas
24	70	9.9	441	1 GUN2_THIEFU	P57674 escherichia
25	70	9.9	753	1 YBHJ_ECOLI	P08506 escherichia
26	69.5	9.8	400	1 DACC_ECOLI	P85837 prochloroth
27	69.5	9.8	661	1 PSAB_PROHO	P93761 capsicum an
28	69	9.7	640	1 INVI_CAVAN	P77570 escherichia
29	68.5	9.6	417	1 YDH_ECOLI	P45623 selaginella
30	68.5	9.6	417	1 HEM2_SEIMA	001837 listeria iv
31	68.5	9.6	864	1 P60_LISIV	002799 sus scrofa
32	68.5	9.6	864	1 STA2_PIG	P97175 mycobacteri
33	68	9.6	159	1 MP63_MYCTU	

ALIGNMENTS

RESULT	ID	Query Match	Length	DB ID	Description
1	DHET_GLUOX	29.7	757	1	005542 gluconobact
AC	005542	29.7	757	1	005542 gluconobact
DT	01-NOV-1997 (Rel. 35, Created)	29.7	757	1	005542 gluconobact
DT	01-NOV-1997 (Rel. 35, Last sequence update)	29.7	757	1	005542 gluconobact
DT	16-OCT-2001 (Rel. 40, Last annotation update)	29.7	757	1	005542 gluconobact
DE	Alcohol dehydrogenase [acceptor] precursor (EC 1.1.99.8) (G3-ADH subunit 1).	29.7	757	1	005542 gluconobact
GN	ADHA.	29.7	757	1	005542 gluconobact
OS	gluconobacter oxydans (gluconobacter suboxydans).	29.7	757	1	005542 gluconobact
OC	Bacteria: Proteobacteria; alpha subdivision: Acetobacteraceae;	29.7	757	1	005542 gluconobact
OC	gluconobacter.	29.7	757	1	005542 gluconobact
OX	NCBI_TaxID=442;	29.7	757	1	005542 gluconobact
RN	[1]	29.7	757	1	005542 gluconobact
RP	SEQUENCE FROM N.A., AND SEQUENCE OF 36-50.	29.7	757	1	005542 gluconobact
RC	STRAIN-IFO 12528;	29.7	757	1	005542 gluconobact
RX	MEDLINE=97208225; PubMed=9055427;	29.7	757	1	005542 gluconobact
RA	Kondo K., Horinouchi S.;	29.7	757	1	005542 gluconobact
RT	*Characterization of the genes encoding the three-component membrane-bound alcohol dehydrogenase from Gluconobacter suboxydans and their expression in Acetobacter pasteurianus.*	29.7	757	1	005542 gluconobact
RT	Appl. Environ. Microbiol. 63:1131-1138(1997).	29.7	757	1	005542 gluconobact
CC	-1- FUNCTION: DEHYDRATION OF PRIMARY ALCOHOLS (EXCEPTION: METHANOL) (BY SIMILARITY).	29.7	757	1	005542 gluconobact
CC	-1- CATALYTIC ACTIVITY: A primary alcohol + acceptor = an aldehyde + reduced acceptor.	29.7	757	1	005542 gluconobact
CC	-1- COFACTOR: POO AND HEME.	29.7	757	1	005542 gluconobact
CC	-1- SUBUNIT: HETEROTRIMER (DEHYDROGENASE, CYTOCHROME AND PROTEIN ADHS), THAT FORMS THE ALCOHOL DEHYDROGENASE COMPLEX.	29.7	757	1	005542 gluconobact
CC	-1- SUBCELLULAR LOCATION: MEMBRANE-BOUND, FACING THE PERIPLASMIC SPACE (POTENTIAL).	29.7	757	1	005542 gluconobact
CC	-1- SIMILARITY: BELONGS TO THE BACTERIAL POO DEHYDROGENASE FAMILY.	29.7	757	1	005542 gluconobact
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@sib-sib.ch).	29.7	757	1	005542 gluconobact
CC	EMBL: D86375; BAA19753.1; -	29.7	757	1	005542 gluconobact
DR	HSSP: 092477; 1FLG.	29.7	757	1	005542 gluconobact
DR	InterPro: IPR001479; Bac_POO.	29.7	757	1	005542 gluconobact
DR	InterPro: IPR002372; Bac_POO_repeat.	29.7	757	1	005542 gluconobact
DR	InterPro: IPR000345; CytC_heme_bind.	29.7	757	1	005542 gluconobact
DR	Pfam: PF01011; Bacterial_POO_6.	29.7	757	1	005542 gluconobact
DR	PROSITE: PS00363; BACTERIAL_POO_1; 1.	29.7	757	1	005542 gluconobact
DR	PROSITE: PS00364; BACTERIAL_POO_2; 1.	29.7	757	1	005542 gluconobact
DR	PROSITE: PS00190; CYTOCHROME_C_1.	29.7	757	1	005542 gluconobact
KW	Oxidoreductase; POO; Heme; Periplasmic; Membrane; Signal.	29.7	757	1	005542 gluconobact
FT	STGNAL	29.7	757	1	005542 gluconobact
FT	CHAIN	29.7	757	1	005542 gluconobact
FT	MOD_RES	29.7	757	1	005542 gluconobact
FT	ACT_SITE	29.7	757	1	005542 gluconobact

FT BINDING 653 653 HEME (COVALENT) (BY SIMILARITY).
 FT BINDING 656 656 HEME (COVALENT) (BY SIMILARITY).
 FT METAL 657 657 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
 SQ SEQUENCE 757 AA; 82968 MW; 39B9F90EB3947581 CRC64.

Query Match 29.7%; Score 211; DB 1; Length 757;
 Best Local Similarity 41.2%; Pred. No. 6.1e-13;
 Matches 56; Conservative 20; Mismatches 52; Indels 8; Gaps 5;

QY 7 LMASGALALIAA--PAFAQV---TPVTDLIANPPAGEMISYGOENYRHSPLQITTT 61
 Db 16 LLSCAAALFAFAVPAVPAEDPTGTAITSSDNGGHP-GDWLSYGRSYSEORYSPLDQINT 74
 QY 62 ENVGOLQIWMARGMOPKQV-TPLIHGVMYLANPGDVIOAIDARTGDLIMEHRRLP- 119
 Db 75 ENVGKLLKLMHYDLDITNGOEGTPIIVGVMTATNMSKMAALDARTGLKLSIDPKVPG 134
 QY 120 NIATLNSFEPTRGMA 135
 Db 135 NIADKGCDDTVNRGAA 150

RESULT 2

DHET_ACEPO STANDARD; PRT; 738 AA.

AC P28036;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Alcohol dehydrogenase [acceptor] precursor (EC 1.1.99.8).
 GN ADHA.
 OS Acetobacter polyoxogenes.
 OC Bacteria; Proteobacteria; alpha subdivision; Acetobacteraceae;
 OC Acetobacter.
 OX NCBI_TaxID=439;
 RN [1]
 RP SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.
 RC STRAIN=NB11028;
 RA MEDLINE=91159482; PubMed=2001402;
 RA Tanaka T., Fukaya M., Takemura H., Takemura K., Okumura H.,
 RA Kawamura Y., Nishiyama M., Horinouchi S., Beppu T.;
 RT Cloning and sequencing of the gene cluster encoding two subunits of
 RT membrane-bound alcohol dehydrogenase from Acetobacter polyoxogenes.";
 RL Biochim. Biophys. Acta 1088:292-300(1991).
 CC -1- FUNCTION: DEHYDRATION OF PRIMARY ALCOHOLS (EXCEPTION: METHANOL).
 CC -1- CATALYTIC ACTIVITY: A primary alcohol + acceptor = an aldehyde +
 CC reduced acceptor.
 CC -1- COFACTOR: POQ AND HEME.
 CC -1- SUBUNIT: HETEROTETRAMER (DEHYDROGENASE, CYTOCHROME, AND TWO
 CC SMALLER UNKNOWN SUBUNITS) THAT FORMS THE ALCOHOL DEHYDROGENASE
 CC COMPLEX.
 CC -1- SUBCELLULAR LOCATION: MEMBRANE-BOUND, FACING THE PERIPLASMIC
 CC SPACE (POTENTIAL).
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POQ DEHYDROGENASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: D00635; BAA00528.1; -
 CC PIR: S14270; S14270.
 CC HSSP: 0924J7; 1EFG.
 CC InterPro: IPR001479; Bac_POQ.
 CC InterPro: IPR002372; Bac_POQ_repeat.
 CC InterPro: IPR000345; CytC_heme_bind.
 CC Pfam: PF01011; Bacterial_POQ_6.
 CC PROSITE: PS00363; BACTERIAL_POQ_1; 1.
 CC PROSITE: PS00364; BACTERIAL_POQ_2; 1.
 CC PROSITE: PS00364; BACTERIAL_POQ_2; 1.

DR PROSITE: PS00190; CYTOCHROME_C; 1.
 KW Oxidoreductase; POQ; Heme; Periplasmic; Membrane; Signal.
 FT SIGNAL 1 35
 FT CHAIN 36 738
 FT BINDING 650 650 ALCOHOL DEHYDROGENASE [ACCEPTOR].
 FT BINDING 653 653 HEME (COVALENT) (BY SIMILARITY).
 FT METAL 654 654 HEME (COVALENT) (BY SIMILARITY).
 FT METAL 654 654 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
 SQ SEQUENCE 738 AA; 80840 MW; 1E2B6ED7BCD92AF6 CRC64.

Query Match 26.3%; Score 187; DB 1; Length 738;
 Best Local Similarity 32.4%; Pred. No. 1.3e-10;
 Matches 44; Conservative 23; Mismatches 59; Indels 10; Gaps 3;

QY 10 SAGLALALAPAFQVTPVDE-----LANPPAGEMISYGOENYRHSPLQITTT 61
 Db 17 TAGTICALISGATVMSADGGGANGEAITHDDHPGMMYGRYSQGRSPDQINR 76
 QY 62 ENVGOLQIWMARGMOPKQV-TPLIHGVMYLANPGDVIOAIDARTGDLIMEHRRLP- 119
 Db 77 SNGVNLKLMYLDLDITNGOEGTPIIVGVMTATNMSKMAALDARTGLKLSIDPKVPG 136
 QY 120 NIATLNSFEPTRGMA 135
 Db 137 NIADKGCDDTVNRGAA 152

RESULT 3

DHET_ACEEU STANDARD; PRT; 739 AA.

AC Q44002; 007952;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Alcohol dehydrogenase [acceptor] precursor (EC 1.1.99.8).
 GN ADH.
 OS Acetobacter europaeus.
 OC Bacteria; Proteobacteria; alpha subdivision; Acetobacteraceae;
 OC Gluconacetobacter.
 OX NCBI_TaxID=33995;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=DES11 / DSM 6160;
 RA Thurner C.A.K.;
 RL Submitted (NOV-1996) to the EMBL/Genbank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: A primary alcohol + acceptor = an aldehyde +
 CC reduced acceptor.
 CC -1- COFACTOR: POQ AND HEME (BY SIMILARITY).
 CC -1- SUBUNIT: HETEROTETRAMER (DEHYDROGENASE, CYTOCHROME, AND TWO
 CC SMALLER UNKNOWN SUBUNITS) THAT FORMS THE ALCOHOL DEHYDROGENASE
 CC COMPLEX (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: MEMBRANE-BOUND, FACING THE PERIPLASMIC
 CC SPACE (POTENTIAL).
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POQ DEHYDROGENASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: X82894; CAA58066.1; -
 CC EMBL: Y09480; CAA70688.1; -
 CC HSSP: 0924J7; 1EFG.
 CC InterPro: IPR001479; Bac_POQ.
 CC InterPro: IPR002372; Bac_POQ_repeat.
 CC InterPro: IPR000345; CytC_heme_bind.
 CC Pfam: PF01011; Bacterial_POQ_6.
 CC PROSITE: PS00363; BACTERIAL_POQ_1; 1.
 CC PROSITE: PS00364; BACTERIAL_POQ_2; 1.
 CC PROSITE: PS00190; CYTOCHROME_C; 1.

[4]
 RN SEQUENCE FROM N.A.
 RC STRAIN-K12 / MG1655;
 RX MEDLINE=97426617; PubMed=9278503;
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Valdes J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of *Escherichia coli* K-12.";
 RL Science 277:1453-1474(1997).
 RN
 RP TOPOLOGY.
 RX MEDLINE=93286127; PubMed=8509415;
 RA Yamada M., Sumi K., Matsushita K., Adachi O., Yamada Y.;
 RT "Topological analysis of quinoprotein glucose dehydrogenase in
 RT *Escherichia coli* and its ubiquinone-binding site.";
 RL J. Biol. Chem. 268:12812-12817(1993).
 RN
 RP 3D-STRUCTURE MODELING.
 RX MEDLINE=96128046; PubMed=8554505;
 RA Cozier G.E., Anthony C.;
 RT "Structure of the quinoprotein glucose dehydrogenase of *Escherichia coli* modelled on that of methanol dehydrogenase from *Methylobacterium extorquens*.";
 RL Biochem. J. 312:679-685(1995).
 CC -1- FUNCTION: GDH IS PROBABLY INVOLVED IN ENERGY CONSERVATION RATHER THAN IN SUGAR METABOLISM.
 CC -1- CATALYTIC ACTIVITY: D-glucose + acceptor = D-glucono-1,5-lactone + reduced acceptor.
 CC -1- COFACTOR: PQQ.
 CC -1- SUBUNIT: MONOMER.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane.
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL PQQ DEHYDROGENASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See [http://www.isb-sib.ch/announce/](http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch) or send an email to license@isb-sib.ch).
 CC
 DR EMBL: X51323; CA35706.1; -;
 DR EMBL: D12651; BAA02174.1; -;
 DR EMBL: D26562; CAB20298.1; -;
 DR EMBL: AE000122; AAC73235.1; -;
 DR PIR: JVO107; JVO107.
 DR HSSP: P38539; 4AAH.
 DR Ecogene; BG10369; gcd.
 DR InterPro: IPR001479; Bac_PQQ.
 DR InterPro: IPR002372; Bac_PQQ_repeat.
 DR Pfam: PF01011; Bacterial_PQQ_7.
 DR PROSITE: PS00363; BACTERIAL_PQQ_1; 1.
 DR PROSITE: PS00364; BACTERIAL_PQQ_2; 1.
 KW Oxidoreductase, PQQ; Transmembrane; Inner membrane; Periplasmic;
 KW Complete proteome.
 FT DOMAIN 1 10
 FT TRANSSEM 11 37 CYTOPLASMIC (PROBABLE).
 FT DOMAIN 38 40 PERIPLASMIC (PROBABLE).
 FT TRANSSEM 41 58 PERIPLASMIC (PROBABLE).
 FT DOMAIN 59 62 CYTOPLASMIC (PROBABLE).
 FT TRANSSEM 63 81 PERIPLASMIC (PROBABLE).
 FT DOMAIN 82 95 PERIPLASMIC (PROBABLE).
 FT TRANSSEM 96 110 CYTOPLASMIC (PROBABLE).
 FT DOMAIN 111 118 CYTOPLASMIC (PROBABLE).
 FT TRANSSEM 119 141 PERIPLASMIC (PROBABLE).
 FT DOMAIN 142 146 PERIPLASMIC (PROBABLE).
 FT ACT_SITE 466 466 BASE (PROBABLE).
 FT CONFLICT 59 59 R -> L (IN REF. 1).
 FT CONFLICT 149 156 TISADATP -> HIKRRCHT (IN REF. 1 AND 2).
 FT CONFLICT 193 193 N -> K (IN REF. 1 AND 2).
 FT CONFLICT 666 666 Q -> H (IN REF. 2).
 SQ SEQUENCE 796 AA; 86747 MW; D9ECC705A12894E9 CRC64;

Query Match 22.3%; Score 158.5; DB 1; Length 796;
 Best Local Similarity 32.0%; Pred. No. 8.4e-08;
 Matches 47; Conservative 22; Mismatches 49; Indels 29; Gaps 7;
 QY 6 LWMAS-----AGALLAAPFAVPTVDLLANPAGEWISVGNQENRRHSPLNQ 58
 DB 135 LTMAGFNDPQELMTISADATPAEA-ISPVADO-----DMPAYGRQEGQRESPLKQ 185
 QY 59 ITTENVGQLQVWA-----RGMQPKV--QVTPLIHDGVMTLANPGVIAIDAKTGD 109
 DB 186 INADNVHNLKEAWFRTGDVKKQPNDEPTNEVTPIKVGDITLCTAHQRFLDALDASGR 245
 QY 110 LIMEHRQDLPNITLNSFGEP-RCMA 135
 DB 246 EKWHYD---PELKTNESEFQHTCRGVS 269
 RESULT 6
 EXAA_PSEAE
 ID EXAA_PSEAE STANDARD; PRT; 623 AA.
 AC Q9Z4J7;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Quinoprotein ethanol dehydrogenase precursor (EC 1.1.99.-) (OEDH).
 GN EXAA OR PA1982.
 OS Pseudomonas aeruginosa.
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 OC Pseudomonas.
 OX NCBI_Taxid=287;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 17933;
 RX MEDLINE=99041560; PubMed=9826187;
 RA Diehl A., Wintzingerode F., Goerlich H.;
 RT "Quinoprotein ethanol dehydrogenase of *Pseudomonas aeruginosa* is a homodimer: sequence of the gene and deduced structural properties of the enzyme.";
 RL Eur. J. Biochem. 257:409-419(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 15692 / PAOI;
 RX MEDLINE=20437337; PubMed=10984043;
 RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P., Hickey M.J., Brinkman F.S.L., Hutnagle W.O., Kowalik D.J., Lagrou M., Gardner R.L., Collier L., Tolentino E., Westbrock-Wadman S., Yuan Y., Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M., Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T., Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
 RT "Complete genome sequence of *Pseudomonas aeruginosa* PAOI, an opportunistic pathogen.";
 RL Nature 406:959-964(2000).
 RN [3]
 RP SEQUENCE OF 1-52 FROM N.A.
 RC STRAIN=ATCC 17933;
 RX MEDLINE=99173751; PubMed=10075429;
 RA Schobert M., Goerlich H.;
 RT "Cytochrome c550 is an essential component of the quinoprotein ethanol oxidation system in *Pseudomonas aeruginosa*: cloning and sequencing of the genes encoding cytochrome c550 and an adjacent acetaldehyde dehydrogenase.";
 RL Microbiology 145:471-481(1999).
 RN [4]
 RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF 35-616.
 RC STRAIN=ATCC 17933;
 RX MEDLINE=20202376; PubMed=10736230;
 RA Ketel T., Diehl A., Knaute T., Stezowski J.J., Hoehne W., Goerlich H.;
 RT "X-ray structure of the quinoprotein ethanol dehydrogenase from *Pseudomonas aeruginosa*: basis of substrate specificity.";
 RL J. Mol. Biol. 297:961-974(2000).

CC -1- FUNCTION: OXIDIZE PRIMARY ALCOHOLS AND ALSO ACTS ON SECONDARY
 CC ALCOHOL, BUT NOT HIGHLY ACTIVE ON METHANOL.
 CC -1- COFACTOR: POQ AND CALCIUM.
 CC -1- SUBUNIT: HOMODIMER.
 CC -1- SUBCELLULAR LOCATION: Periplasmic.
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POQ DEHYDROGENASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 CC EMBL: AJ009858; CAA08896.1; -
 CC EMBL: AF004624; AAG05370.1; -
 CC EMBL: AF068264; AAC79657.1; -
 CC PDB: 1FUG; 30-AUG-00.
 CC InterPro: IPR001479; Bac_POQ.
 CC InterPro: IPR002372; Bac_POQ_repeat.
 CC Pfam: PF01011; Bacterial_POQ_7.
 CC PROSITE: PS00363; BACTERIAL_POQ_1; FALSE_NEG.
 CC PROSITE: PS00364; BACTERIAL_POQ_2; FALSE_NEG.
 CC Oxidoreductase; POQ; Periplasmic; Signal; Calcium; 3D-structure;
 KW Complete proteome.
 KW SIGNAL 1 34
 FT CHAIN 35 623 QUINOPROTEIN ETHANOL DEHYDROGENASE.
 FT DISULFID 139 140
 SO SEQUENCE 623 AA; 68123 MW; 32DDE5DF20B291D6 CRC64;
 Query Match 19.3%; Score 137; DB 1; Length 623;
 Best Local Similarity 31.3%; Pred. No. 7.9e-06;
 Matches 40; Conservative 20; Mismatches 57; Indels 10; Gaps 4;
 QY 3 PTSLMASAGALAL---LAAPAFQVTPVDELLAN--PPAGEWISYGQDENVRSPLT 57
 DB 9 PAGLRPSLHCLARVALAGSAGALADYWDIANDDKTGDVLYQGMGTHQRMSPLE 68
 QY 58 QITENNGQGLQVWARGN---QPKVQVPLIHGVMYLANPGDVIAIDAKTGDLIWE 113
 DB 69 QVNDNVKELIPAMVSYSRSGDEKQGO--ESQAVISDGVIVYASYSRLFALDAKTGRKRLMT 127
 QY 114 HRRQLPN 120
 DB 128 YNHRLLP 134
 RESULT 7
 DHG_GLUOX STANDARD: PRT: 808 AA.
 AC P27175;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Glucose dehydrogenase [pyrroloquinoline-quinone] precursor
 DE (EC 1.1.99.17).
 GN GDH.
 OS Gluconobacter oxydans (Gluconobacter suboxydans).
 OC Bacteria; Proteobacteria; alpha subdivision; Acetobacteraceae;
 OC Gluconobacter.
 OC NCBI_TaxID=442;
 RN [1]
 RN SEQUENCE FROM N.A.
 RP MEDLINE=92017653; PubMed=1833518;
 RX Cleton-Jansen A.-M., Dekker S., van de Putte P., Goosen N.;
 RT "A single amino acid substitution changes the substrate specificity
 RT of quinoprotein glucose dehydrogenase in Gluconobacter oxydans.";
 RL Mol. Gen. Genet. 229:206-212(1991).
 RN [2]
 RN STRAIN=LMD 79.41;
 RP SEQUENCE FROM N.A.
 RP REVISION TO 213.
 RA Goosen N.;

RL Submitted (FEB-1992) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: D-glucose + acceptor = D-glucono-1,5-lactone +
 CC reduced acceptor.
 CC -1- COFACTOR: POQ.
 CC -1- SUBCELLULAR LOCATION: OUTSIDE OF THE PERIPLASMIC MEMBRANE.
 CC -1- MISCELLANEOUS: THE P1 FORM CAN OXIDIZE ONLY D-GLUCOSE, WHILE THE
 CC P2 FORM CAN ALSO OXIDIZE MALTOSE. THE SEQUENCE OF P1 FORM IS SHOWN
 CC HERE.
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POQ DEHYDROGENASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 CC EMBL: X62710; CAA44594.1; ALT_SEQ.
 CC PIR: S17716; QPKEX.
 CC InterPro: IPR001479; Bac_POQ.
 CC InterPro: IPR002372; Bac_POQ_repeat.
 CC Pfam: PF01011; Bacterial_POQ_7.
 CC PROSITE: PS00363; BACTERIAL_POQ_1; 1.
 CC PROSITE: PS00364; BACTERIAL_POQ_2; 1.
 CC Oxidoreductase; POQ; Transmembrane; Periplasmic; Signal.
 KW SIGNAL 1 33
 FT CHAIN 34 808
 FT TRANSMEM 35 54
 FT TRANSMEM 59 76
 FT TRANSMEM 94 108
 FT TRANSMEM 123 138
 FT ACT SITE 470 470
 FT VARIANT 788 788 H -> N (IN P2 FORM).
 FT SEQUENCE 808 AA; 87567 MW; 0F416DDA78652445 CRC64;
 Query Match 18.9%; Score 134; DB 1; Length 808;
 Best Local Similarity 27.8%; Pred. No. 2.1e-05;
 Matches 35; Conservative 20; Mismatches 51; Indels 20; Gaps 3;
 QY 13 ALALAAPAFQVTPVDELLA-----NPPAGEWISYGQDENVRSPLQIQT 61
 DB 128 AVLALFASLFTDPHDISGELPTQIANASPADPDNVASEHAGRTQAGDRKSPLOQINA 187
 QY 62 ENYGQGLQVW-----ARCMQPKV--QVPLIHGVMYLANPGDVIAIDAKTGDLIWM 112
 DB 188 TNSVSNLKVAMHHTKDMANSNDGEGQTNENATPIEFNNITLYMCSLHQKLFAYDQATGNVAKM 247
 QY 113 EHRQL 118
 DB 248 VYDPKL 253
 RESULT 8
 DHGA_ACICA STANDARD: PRT: 801 AA.
 ID DHGA_ACICA
 AC P05465;
 DT 01-NOV-1988 (Rel. 09, Created)
 DT 01-NOV-1988 (Rel. 09, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Glucose dehydrogenase-A [pyrroloquinoline-quinone] precursor
 DE (EC 1.1.99.17) (Quinoprotein glucose DH) (GDH-A).
 GN GDHA.
 OS Acinetobacter calcoaceticus.
 OC Bacteria; Proteobacteria; gamma subdivision; Moraxellaceae;
 OC Acinetobacter.
 OC NCBI_TaxID=471;
 RN [1]
 RN SEQUENCE FROM N.A.
 RP STRAIN=LMD 79.41;
 RC MEDLINE=88289368; PubMed=3399393;
 RX

RA Clifton-Jansen A.-M., Goosen N., Odle G., van de Putte P.;
 RT "Nucleotide sequence of the gene coding for quinoprotein glucose
 dehydrogenase from *Acinetobacter calcoaceticus*.";
 RL Nucleic Acids Res. 16:6228-6228(1988).
 CC -1- CATALYTIC ACTIVITY: D-glucose + acceptor = D-glucono-1,5-lactone +
 CC reduced acceptor.
 CC -1- COFACTOR: POQ.
 CC -1- SUBUNIT: MONOMER.
 CC -1- SUBCELLULAR LOCATION: Periplasmic.
 CC -1- MISCELLANEOUS: ACINETOBACTER CALCOCOETICUS CONTAINS TWO DIFFERENT
 CC POQ DEPENDENT ENZYMES WITH GDH ACTIVITY. GDH-A PREFERENCES
 CC 2-DEOXYGLUCOSE AS SUBSTRATE, THE SPECIFIC SUBSTRATES FOR GDH-B ARE
 CC DISACCHARIDES.
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POQ DEHYDROGENASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: X07235; CAA30222.1; -
 CC PIR: S00943; S00943.
 DR InterPro: IPR001479; Bac_POQ.
 DR InterPro: IPR002372; Bac_POQ_repeat.
 DR Pfam: PF01011; Bacterial_POQ_7.
 DR PROSITE: PS00363; BACTERIAL_POQ_1; 1.
 DR PROSITE: PS00364; BACTERIAL_POQ_2; 1.
 KM Oxidoreductase; POQ; Transmembrane; Periplasmic; Signal.
 FT SIGNAL 1 33
 FT CHAIN 34 801
 FT TRANSMEM 39 55
 FT TRANSMEM 59 79
 FT TRANSMEM 94 108
 FT TRANSMEM 119 138
 FT ACT_SITE 471 471
 SQ SEQUENCE 801 AA; 86956 MW; 2F4EA222EB5450D8 CRC64;

Query Match 17.3%; Score 123; DB 1; Length 801;
 Best Local Similarity 30.7%; Pred. No. 0.00024;
 Matches 39; Conservative 16; Mismatches 48; Indels 24; Gaps 5;

QY 20 PAFAYPTVDELPPAGEMISYGNONEYRHSPLTOTTENVGOLQVWARGQPK 79
 DB 152 PETAQAVPVAAE-----SDWPAYGRTOAGVAYSPKQINDQNVADLAVMT--LRIGD 202
 QY 80 V-----QVPLIHGQVYLANPCGVIAIDAKTGDLIWEHRRQIPNATLNSFG 128
 DB 203 LKTDNDSEGTNQVTPKIGNNFICTAHQOLIAIDPATGKEKW--RDPKLTDKSPQ 259
 QY 129 EPT-RGM 134
 DB 260 HLTGRCV 266

RESULT 9
 DHM1_PARDE STANDARD; PRT; 631 AA.
 AC P12293;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Methanol dehydrogenase subunit 1 precursor (EC 1.1.99.8) (MDH large
 DE alpha subunit) (MEDH).
 GN MOXF.
 OS Paracoccus denitrificans.
 OC Bacteria; Proteobacteria; alpha subdivision; Rhodobacter group;
 OC Paracoccus.
 ON NCB1_TaxID=266;

RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 33-49.
 RX MEDLINE=87307969; PubMed=3114231;
 RA Harms N., de Vries G.E., Maurer K., Hoogenbijk J., Stoutamer A.H.;
 RT "Isolation and nucleotide sequence of the methanol dehydrogenase
 RT structural gene from *Paracoccus denitrificans*.";
 RL J. Bacteriol. 169:3969-3975(1987).
 CC -1- CATALYTIC ACTIVITY: A primary alcohol + acceptor = an aldehyde +
 CC reduced acceptor.
 CC -1- COFACTOR: POQ.
 CC -1- SUBCELLULAR LOCATION: Periplasmic.
 CC -1- MISCELLANEOUS: MDH IS THE MAJOR PROTEIN IN THE CELL DURING GROWTH
 CC ON METHANOL (IN P. DENITRIFICANS MDH CONSTITUTES UP TO 15% OF THE
 CC TOTAL CELL PROTEIN).
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POQ DEHYDROGENASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: M1739; AAA8366.1; -
 CC HSSP: P38539; 4AAH.
 DR InterPro: IPR001479; Bac_POQ.
 DR InterPro: IPR002372; Bac_POQ_repeat.
 DR Pfam: PF01011; Bacterial_POQ_7.
 DR PROSITE: PS00363; BACTERIAL_POQ_1; 1.
 DR PROSITE: PS00364; BACTERIAL_POQ_2; 1.
 KM Oxidoreductase; POQ; Signal; Methanol utilization; Periplasmic.
 FT SIGNAL 1 32
 FT CHAIN 33 631
 FT DISULFID 135 136
 FT DISULFID 418 447
 FT ACT_SITE 335 335
 SQ SEQUENCE 631 AA; 69799 MW; 0934DC93FFC5730B CRC64;

Query Match 16.8%; Score 119.5; DB 1; Length 631;
 Best Local Similarity 32.18%; Pred. No. 0.00041;
 Matches 45; Conservative 26; Mismatches 54; Indels 15; Gaps 9;

QY 4 TSLWASAGALALL-AAFAQVPTDEL--LANPAGEMISYGNONEYRHSPLQIT 60
 DB 12 SSLAAVAVMGLAVLTATATA-----NDQVLELAKDRA-NWMTGRQDNQNTSEM7DIN 65
 QY 61 TENVGOLQVW--ARGMQPKVQVTPPLIHGQVYLANP-GDVIAIDA-KTGDLIWEHR- 115
 DB 66 KENVKQLRPAMSESTGVLDGH-EGPILVVGDRMFTHFPFPTPALDLNFGKILMQNPK 124
 QY 116 RQLEPIATLNSFGEPTRGMA 135
 DB 125 KQNPARTVACCDVYNRGLA 144

RESULT 10
 DHM1_METOR STANDARD; PRT; 626 AA.
 AC P15279;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-APR-1990 (Rel. 14, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Methanol dehydrogenase subunit 1 precursor (EC 1.1.99.8) (MDH large
 DE alpha subunit) (MEBH).
 GN MOXF.
 OS Methylobacterium organophilum XX.
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Methylobacterium group; Methylobacterium.
 ON NCB1_TaxID=410;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 29-43.

DB 67 LRPANFSTGLNGH-EGAPLVGDKMYHTSPNNTFALGLDDPGTILMOKFKONPAA 125
 QY 122 ATLNFGEPTRGMA 135
 DB 126 RAVACCDLVNRGLA 139

RESULT 12
 ID XOXF_PARDE STANDARD; PRT; 600 AA.
 AC P29968;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Putative dehydrogenase XOXF precursor (EC 1.1.99.-).
 GN XOXF.
 OS Paracoccus denitrificans.
 OC Bacteria; Proteobacteria; alpha subdivision; Rhodobacter group;
 CC Paracoccus.
 OX NCBI_Taxid=266;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Harms N.;
 RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
 RP SEQUENCE OF 380-600 FROM N.A.
 RC STRAIN=PD 1235;
 RX MEDLINE=92041583; PubMed=1657873;
 RA Ras J., Reijnders W.N.M., van Spanning R.J.M., Harms N., Oltmann L.F.,
 RA Stouthamer A.H.;
 RT "Isolation, sequencing, and mutagenesis of the gene encoding
 RT cytochrome c551 of Paracoccus denitrificans and characterization of
 RT the mutant strain."
 RL J. Bacteriol. 173:6971-6979(1991).
 CC -1- COFACTOR: POQ (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POQ DEHYDROGENASE FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: U34346; AAC44555.1;
 DR EMBL: M75583; AAA25574.1;
 DR PIR: A41378; A41378.
 DR HSSP: P38539; 4AAH.
 DR InterPro: IPR002372; Bac_POQ_repeat.
 DR Pfam: PF01011; Bacterial_POQ; 7.
 KW Oxidoreductase; POQ; Signal.
 FT SIGNAL 1 21
 FT CHAIN 22 600
 FT ACT_SITE 318 318
 FT SEQUENCE 600 AA; 65159 MW; DCA96F1BC5A3CE CRC64;
 FT PUTATIVE DEHYDROGENASE XOXF.
 FT BASE (POTENTIAL).
 FT SEQUENCE 600 AA; 65159 MW; DCA96F1BC5A3CE CRC64;

Query Match 13.9%; Score 99; DB 1; Length 600;
 Best Local Similarity 27.5%; Pred. No. 0.038;
 Matches 38; Conservative 25; Mismatches 53; Indels 22; Gaps 7;

QY 10 SAGALALLAAPFAOVPTPTDELAMP-----PAGEWISTGONCEYRHSPLTQITTE 62
 DB 6 NAACIALINSGTAA-----LANEORAGRDROAPQWALOMGDYANTRYSTLDQINKMD 56
 QY 63 NGGOLQIVM--ARGMOPGVQVYPTLIHGGVWYLANP-GDVQIAIDAK-TGDLWEHR-RQ 117
 DB 57 NKDLRVAMTSTGYLRGH-EGSPLYIGDVIVHTPPFRVFAIDLNDNGKILMYEPOQ 115
 QY 118 LFNIALTNSFGEPTRGMA 135
 DB 116 DENVIAMVMSCTVYRGLS 133

RESULT 13
 ID QUITA_ACICA STANDARD; PRT; 809 AA.
 AC Q59086;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Quinate/shikimate dehydrogenase [pyrroloquinoline-quinone]
 DE (EC 1.1.99.25) (NAD(P)-independent quinate dehydrogenase).
 GN QUITA.
 OS Acinetobacter calcoaceticus.
 OC Bacteria; Proteobacteria; gamma subdivision; Moraxellaceae;
 CC Acinetobacter.
 OX NCBI_Taxid=471;
 RN [1]
 RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
 RC STRAIN=BD413 / ADP1;
 RX MEDLINE=95095936; PubMed=8002591;
 RA Elsomore D.A., Ornstion L.N.;
 RT "The pca-pob supraoperonic cluster of Acinetobacter calcoaceticus
 RT contains quita, the structural gene for quinate-shikimate
 RT dehydrogenase."
 RL J. Bacteriol. 176:7659-7666(1994).
 RP SEQUENCE OF 1-20 FROM N.A.
 RC STRAIN=BD413 / ADP1.
 RX MEDLINE=96011389; PubMed=7592351;
 RA Elsomore D.A., Ornstion L.N.;
 RT "Unusual ancestry of dehydratases associated with quinate catabolism
 RT in Acinetobacter calcoaceticus."
 RL J. Bacteriol. 177:5971-5978(1995).
 CC -1- FUNCTION: CAN ACT EITHER ON QUINATE OR ON SHIKIMATE.
 CC -1- CATALYTIC ACTIVITY: quinate + pyrroloquinoline-quinone = 5-
 CC -1- DEHYDROSHIKIMATE + reduced pyrroloquinoline-quinone.
 CC -1- CATALYTIC ACTIVITY: SHIKIMATE + pyrroloquinoline-quinone = 3-
 CC -1- DEHYDROSHIKIMATE + REDUCED PYRROLOQUINOLINE-QUINONE.
 CC -1- PATHWAY: QUINIC ACID CATABOLIC PATHWAY. FIRST STEP. THIS PATHWAY
 CC ALLOWS GROWTH OF BACTERIA WITH QUINATE BY ITS CONVERSION TO
 CC PROTOCATECHUATE AND SUBSEQUENT METABOLISM BY THE BETA-KETOADIPATE
 CC PATHWAY.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -1- INDUCTION: BY PROTOCATECHUATE.
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POQ DEHYDROGENASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: L05770; AAC37161.1;
 DR InterPro: IPR001479; Bac_POQ.
 DR InterPro: IPR002372; Bac_POQ_repeat.
 DR Pfam: PF01011; Bacterial_POQ; 7.
 DR PROSITE: PS00363; BACTERIAL_POQ_1; FALSE_NEG.
 DR PROSITE: PS00364; BACTERIAL_POQ_2; FALSE_NEG.
 KW Oxidoreductase; POQ; Quinate metabolism; Transmembrane.
 FT OXIDOREDUCTASE; POQ 34
 FT TRANSMEM 41 61
 FT TRANSMEM 68 88
 FT TRANSMEM 90 110
 FT TRANSMEM 127 147
 FT SEQUENCE 809 AA; 88196 MW; 71F67CBEA62BFCB CRC64;

Query Match 13.9%; Score 99; DB 1; Length 809;
 Best Local Similarity 29.0%; Pred. No. 0.053;
 Matches 42; Conservative 15; Mismatches 54; Indels 34; Gaps 6;

Fri May 24 11:27:20 2002

us-08-934-506a-5_copy_1_135.rsp

Page 10

Search completed: May 24, 2002, 10:15:50
Job time: 768 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2002, 10:17:31 ; Search time 93.93 seconds

(without alignments)
248.635 Million cell updates/sec

Title: US-08-934-506a-5_COPY_1_135

Sequence: 1 MKPTSLMAGALALAP.....RQENITATINSFGEPTGMA 135

Scoring table:

Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Listing first 45 summaries

```

1: SPTREMBL_19:*
2: sp.archaea:*
3: sp.bacteria:*
4: sp.fungi:*
5: sp.human:*
6: sp.invertebrate:*
7: sp.mammal:*
8: sp.mhc:*
9: sp.organelle:*
10: sp.phage:*
11: sp.plant:*
12: sp.podent:*
13: sp.virus:*
14: sp.vertibrate:*
15: sp.unclassified:*
16: sp.rvirts:*
17: sp.bacteriap:*
18: sp.archaeap:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	268	37.7	608	2	Q93RE9 pseudogluc
2	191	26.9	742	2	O53362 acetobacter
3	178.5	25.1	691	2	O9AP95 pseudomonas
4	174	24.5	623	2	O9AGM3 pseudomonas
5	168.5	23.7	698	2	O9KH03 alcaligenes
6	167	23.5	786	2	P95466 pantoea cit
7	165.5	23.3	708	2	O46444 comamonas t
8	156	22.0	695	2	O9F9U2 pseudomonas
9	155	21.8	785	16	O98KF6 rhizobium m
10	146.5	20.6	777	16	O92RB3 rhizobium m
11	143	20.1	801	2	O9EYW8 rhizobium m
12	140.5	19.8	601	16	O9EYW8 rhizobium m
13	140.5	19.8	601	16	O9EYW8 rhizobium m
14	121	17.0	790	2	O9X2S5 pantoea cit
15	119	16.8	601	2	P71509 methyllobact
16	118.5	16.7	695	2	O934G0 pseudomonas

17	111.5	15.7	629	2	O9A048 methylobact
18	105.5	14.9	633	2	O24759 hyphomicrob
19	93.5	13.2	639	2	P77931 pseudomonas
20	93.5	13.2	644	2	O52551 pseudomonas
21	89	12.5	499	2	O51842 thiodacillin
22	86	12.1	613	17	O28873 archaeoglob
23	84.5	11.9	599	2	O9L935 methylovoru
24	83.5	11.8	2832	2	O93N05 agrobacteri
25	82.5	11.6	485	17	O96XP7 sulfolobus
26	82	11.5	668	16	O9RRH3 phallusia m
27	81	11.4	680	5	O9N9G7 pseudomonas
28	80.5	11.3	647	16	O9HUB1 pseudomonas
29	79.5	11.2	846	2	O86092 mycobacteri
30	78.5	11.1	509	16	P96368 listeria in
31	78	11.0	207	16	O92D89 listeria in
32	76	10.7	306	16	O9RY72 delnococcus
33	76	10.7	386	16	O9KTW8 vibrio chol
34	76	10.7	573	2	O59540 methylophil
35	76	10.7	4809	2	O93HH0 streptomyce
36	75.5	10.6	407	16	O9PG38 xylella fas
37	75.5	10.6	684	2	O9S282 streptomyce
38	75.5	10.6	1419	5	O9VYL3 P73139 synchocyst
39	75.5	10.6	3972	16	P73139 O9CK11 pasteurilla
40	75	10.6	183	16	O9CKL1 pasteurilla
41	74.5	10.5	969	16	O9RZT5 O9K115 streptomyce
42	74	10.4	294	2	O50152 streptomyce
43	74	10.4	380	16	O9HXJ7 pseudomonas
44	74	10.4	525	16	O93IE9 O93169 staphylococ
45	74	10.4	755	2	O9WMA1 O9WMA1 agrobacteri

ALIGNMENTS

```

RESULT 1
ID 093RE9 PRELIMINARY: PRT: 608 AA.
AC 093RE9;
AT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)
DE ALCOHOL DEHYDROGENASE.
GN ADH.
OS Pseudoglucobacter saccharoketogenes.
OC Bacteria; Pseudoglucobacter.
NCBI_Taxid=133921;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IFO 14464;
RT Shitaba T., Saito Y.;
RL "Alcohol dehydrogenase."
DR EMBL: AB046580; BAB62258.1;
SQ SEQUENCE 608 AA; 65101 MW; 0ACBC97AE11BA570 CRC64;

```

Query Match 37.7%; Score 268; DB 2; Length 608;
Best Local Similarity 40.4%; Pred. No. 7.3e-18;
Matches 59; Conservative 21; Mismatches 46; Indels 20; Gaps 2;

QY	10	SAGALALAPAFAPAO-----VTPVDELLANPPAGEWISYGOEN 50
DB	16	STALIASLSPAFQAHANNAAPESKAGSAIENFQVTFADDLAKRNPANPILRGYOG 75
QY	51	YRHSPLQITTEENVGOLQLYWARGMGKQVPLIHDGYMYLANPGVDYQADAKTGDL 110
DB	76	WGSPPLDQINKNDVGDOLVWSRTMEPSNEGAIAIVNGVIFGNTNDVLOALDKGTSL 135
QY	111	IWEHROLPNLAT-LNSFGEPTGMA 135
DB	136	IWEHRLKLPASKRINSLGAAKRSIA 161

```

RESULT 2
Q53362 ID 053362 PRELIMINARY; PRT; 742 AA.
AC 053362: 044159:
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DE DEHYDROGENASE SUBUNIT OF ALCOHOL DEHYDROGENASE.
OS Acetobacter pasteurianus (Acetobacter turbidans).
OC Bacteria: Proteobacteria; alpha subdivision; Acetobacteraceae;
OX NCBI_TaxID=438;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NC11380;
RX MEDLINE=94042848; PubMed=8226628;
RA Takemura H., Kondo K., Horinouchi S., Beppu T.;
RT "Induction by ethanol of alcohol dehydrogenase activity in Acetobacter
  pasteurianus."
RL J. Bacteriol. 175:6857-6866(1993).
DR EMBL; D13893; BAA40252.1; -.
DR HSSP; Q924J7; 1F1G.
DR InterPro; IPR001479; Bac_P00.
DR InterPro; IPR002372; Bac_P00_repeat.
DR Pfam; PF01011; Bacterial_P00; 6.
DR PROSITE; PS00363; BACTERIAL_P00_1; 1.
DR PROSITE; PS00364; BACTERIAL_P00_2; 1.
SQ SEQUENCE 742 AA; 81680 MW; 031C4E9B6B13EC98 CRC64;

Query Match
Best Local Similarity 26.9%; Score 191; DB 2; Length 742;
Matches 47; Conservative 18; Mismatches 58; Indels 4; Gaps 3;

QY 11 AGALALLAPAFQVPTVDLDEL--ANPPAGEWISYQNGQENYRHSPLTQITTEVNGOQ 68
DB 23 AALPYAAVADPARADGGGNGTGEIITHDDHPENMLSTGRYSRSPLOQINRSVNGDLK 82
QY 69 LVWARGMOGKQV--VPLIHGCVMTLANPGDVIQAIIDAKTGLDIMEHRRLP--NTATLNS 126
DB 83 LAWYTLDTNRGQATPLVVDGIMVATNWSKMEALDAATGKLMLQYDPKVGNIADKGC 142
QY 127 FGEPTRG 133
DB 143 CQTVNGC 149

RESULT 3
Q9AF95 ID 09AF95 PRELIMINARY; PRT; 691 AA.
AC 09AF95:
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DE 1-BUTANOL DEHYDROGENASE BDH.
OS Pseudomonas butanovora.
OC Bacteria: Proteobacteria; gamma subdivision; Pseudomonadaceae;
OX NCBI_TaxID=86174;
RN [1]
RP SEQUENCE FROM N.A.
RA Vangnai A.S., Arp D.J., Sayavedra-Soto L.A.;
RT "Characterization of the expression of two distinct alcohol
  dehydrogenases involved in butane metabolism in Pseudomonas
  butanovora."
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF355798; AAK27220.2; -.
SQ SEQUENCE 691 AA; 75070 MW; 4FC7FD20CDA14E64 CRC64;

Query Match
Best Local Similarity 25.1%; Score 178.5; DB 2; Length 691;
Matches 31; Conservative 22; Mismatches 28; Indels 1; Gaps 1;

```

```

QY 39 GEMISYQNGQENYRHSPLTQITTEVNGOQLVWARGMOGPK-VQYTPLIHGVMTLANPG 97
DB 41 GEMRTGYDAGTRYSPLQITPDNAKELGLVMSYDLESRCVATPTIVDGVMTVTAAPW 100
QY 98 DVIQATIDAKTGLDIMEHRRLP 119
DB 101 SVVHALDVNSGRKMLMTPEVP 122

RESULT 4
Q9AGW3 ID 09AGW3 PRELIMINARY; PRT; 623 AA.
AC 09AGW3:
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DE 1-BUTANOL DEHYDROGENASE BOH PRECURSOR.
OS Pseudomonas butanovora.
OC Bacteria: Proteobacteria; gamma subdivision; Pseudomonadaceae;
OX NCBI_TaxID=86174;
RN [1]
RP SEQUENCE FROM N.A.
RA Sayavedra-Soto L.A., Vangnai A.S., Arp D.J.;
RT "Characterization of the expression of two distinct alcohol
  dehydrogenases involved in butane metabolism in Pseudomonas
  butanovora."
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF326086; AAK15506.1; -.
DR HSSP; Q924J7; 1F1G.
DR InterPro; IPR002372; Bac_P00_repeat.
DR Pfam; PF01011; Bacterial_P00; 4.
KW Signal.
SQ SEQUENCE 623 AA; 67553 MW; A573A1C64AC12D55 CRC64;

Query Match
Best Local Similarity 24.5%; Score 174; DB 2; Length 623;
Matches 49; Conservative 18; Mismatches 46; Indels 16; Gaps 5;

QY 2 KPTSLWASAGALLAPAFQVPTVDLDEL--PPAGEWISYQNGQENYRHSPLTQI 59
DB 7 KPFALRAIVATAAALSLPA-AAVTDVTMEDIANDKHTGDVLTGGLAAGHSPKAI 65
QY 60 TTEVNGOQLVMA-----RGMOGKQVYTPLIHGVMTLANPGDVIQAIIDAKTGLDI 111
DB 66 NTDVNAVLVPAWSFSGEKORGE--AAV--LVHGVIVATASYSRIFAIDARSGKRL 120
QY 112 WEHRRLPN 120
DB 121 WEYNAKLPD 129

RESULT 5
Q9KH03 ID 09KH03 PRELIMINARY; PRT; 698 AA.
AC 09KH03:
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DE 1-BUTANOL DEHYDROGENASE.
OS Alcaligenes eutrophus (Ralstonia eutropha).
OC Bacteria: Proteobacteria; beta subdivision; Ralstonia group;
OX NCBI_TaxID=510;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21125557; PubMed=11222593;
RA Zarrin G., Schrader T., Andreassen J.R.;
RT "Catalytic and Molecular Properties of the Quinolomoprotein
  tetrahydrofuryl Alcohol Dehydrogenase from Ralstonia eutropha

```


CC - FUNCTION: CATALYZES THE DYE-LINKED OXIDATION OF PRIMARY ALCOHOLS
 CC TO THE CORRESPONDING ALDEHYDES AND THE (SUBSEQUENT) OXIDATION OF
 CC THE ALDEHYDES TO CARBOXYLIC ACIDS.
 CC - COFACTOR: POO, HEME, AND CALCIUM.
 CC - SUBUNIT: MONOMER.
 CC - SUBCELLULAR LOCATION: PERIPLASMIC (POTENTIAL).
 CC - INDUCTION: BY ETHANOL AND BUTANOL.
 CC - SIMILARITY: TO ALCOHOL DEHYDROGENASE (ACCEPTOR) PROTEINS.
 DR EMBL: X81880; CAA57464.1; -
 DR HSSP: Q924J7; 1FLG.
 DR InterPro: IPR002372; Bac_POO_repeat.
 DR InterPro: IPR003088; Cyt_C1.
 DR InterPro: IPR002329; Cyt_C1C.
 DR Pfam: PF01011; Bacterial_POO; 6.
 DR Pfam: PF00034; Cytochrome_c; 1.
 DR PRINTS: PR00605; CYTOCHROME_C.
 KW Signal: POO; Heme; Calcium; Oxidoreductase; Periplasmic.
 FT SIGNAL 1 31
 FT CHAIN 32 708
 FT BINDING 635 635
 FT BINDING 638 638
 FT METAL 639 639
 SO SEQUENCE 708 AA; 76822 MW; 99AB54BDD6ACCB3 CRC64;

Query Match 23.3%; Score 165.5; DB 2; Length 708;
 Best Local Similarity 31.8%; Pred. No. 9.3e-08;
 Matches 41; Conservative 23; Mismatches 50; Indels 15; Gaps 4;

OY 3 PTSLMSAGALLAAPAFQVTPVY-----DELANPP-AGEWISYGOQEN 50
 DB 12 PGRWMLAACLG--SAAAFQGTGPAQAAAVORVDGDFIRMAARTPMPRTGVYAE 69
 OY 51 YRHSPLQITTEYNGQOLVWARGMPGK-VQVPLIHDPVMTLANGDVIOADAKTGD 109
 DB 70 TRYSLRDLQINMANVKDGLMSVYLESTREVENTPVVDIMYVSNMVAHIDRTGN 129
 OY 110 LIWEHRQL 118
 DB 130 RIWTYDPOI 138
 RESULT 8
 O9F9U2 PRELIMINARY; PRT; 695 AA.
 AC O9F9U2;
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
 DE ALCOHOL DEHYDROGENASE.
 GN ADH.
 OS Pseudomonas stutzeri (Pseudomonas perfectomarina).
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 OC Pseudomonas.
 OX NCBI_TaxID=316;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BC-2;
 RA Chang C.-H., Herrick J.B., Okinaka R.T., Brainerd J.B.,
 RA Terwilliger T.C.;
 RT Identification and characterization of genes activated by 2-
 RT chloroethanol in Pseudomonas stutzeri BC-2;
 RT Submitted (Aug-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF176640; AAG09249.1; -
 DR HSSP: Q924J7; 1FLG.
 DR InterPro: IPR002372; Bac_POO_repeat.
 DR InterPro: IPR003088; Cyt_C1.
 DR InterPro: IPR002329; Cyt_C1C.
 DR Pfam: PF01011; Bacterial_POO; 6.
 DR Pfam: PF00034; Cytochrome_c; 1.
 DR PROSITE: PS00190; CYTOCHROME_C; UNKNOWN_1.
 DR PROSITE: PS00190; CYTOCHROME_C; UNKNOWN_1.
 SO SEQUENCE 695 AA; 75842 MW; D29A698A642E1A83 CRC64;

Query Match 22.0%; Score 156; DB 2; Length 695;
 Best Local Similarity 33.6%; Pred. No. 7.7e-07;
 Matches 36; Conservative 22; Mismatches 45; Indels 4; Gaps 3;

OY 15 ALAAPAFQVTPVTELL--ANPAGWISYGOQENYRHSPLQITTEYNGQOLVWA 72
 DB 16 ALLVA-AGAAVAYDEAATASQDSSEWLSHRTYAEQFSPKQIDAGNKGILAWY 74
 OY 73 RGMOPGK-VQVPLIHDPVMTLANGDVIOADAKTGDILWEHRQL 118
 DB 75 LDLENNRGLFATPLVSDGVLYASTLSMSRVAVYDLSRGKRLMOPDPQY 121

RESULT 9
 O98RF6 PRELIMINARY; PRT; 785 AA.
 AC O98RF6;
 DT 01-OCT-2001 (TrEMBLrel. 18, Created)
 DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
 DT 01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
 DE GLUCOSE DEHYDROGENASE.
 GN MLL1500.
 OS Rhizobium loti (Mesorhizobium loti).
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Phyllobacteriaceae; Mesorhizobium.
 OX NCBI_TaxID=381;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MAFP303099;
 RC MEDLINE=21082930; PubMed=11214968;
 RA Kaneko T., Nakamura Y., Sato S., Asanizu E., Kato T., Sasamoto S.,
 RA Matanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
 RA Kishida Y., Kiyokawa C., Kohara M., Matsuno M., Matsuno A.,
 RA Mochizuki Y., Nakayama S., Nakazaki N., Shimo S., Sugimoto M.,
 RA Takeuchi C., Yamada M., Tabata S.;
 RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
 RT Mesorhizobium loti";
 RT DNA Res. 7:331-338(2000).
 RL EMBL: AP002997; BAB48856.1;
 DR InterPro: IPR002372; Bac_POO_repeat.
 DR Pfam: PF01011; Bacterial_POO; 7.
 KW Complete proteome.
 SO SEQUENCE 785 AA; 84083 MW; IDC224B7A1C25470 CRC64;

Query Match 21.8%; Score 155; DB 16; Length 785;
 Best Local Similarity 31.7%; Pred. No. 1.1e-06;
 Matches 46; Conservative 13; Mismatches 46; Indels 40; Gaps 4;

OY 8 WASAG-----ALAAPAFQVTPVY-----ELANPAGE 40
 DB 111 WARRGLAGPDGRAPILLVILVLSAVGYSMTADPKGIDGALDTOKVIPKANLGVDPAGE 170
 OY 41 WISYGOQENYRHSPLQITTEYNGQOLVWARGMPGK-----QVPLIHG 89
 DB 171 WHYGRTOFGORYSPLDITPDNVANLQPAWT--YRIGDVAGPDDIGETTYQVPLIAGD 228
 OY 90 WMYLANPDVIOADAKTGDILWEH 114
 DB 229 TLYICTPHNFAIVDAVGAAGKEKRWY 253

RESULT 10
 O92RB3 PRELIMINARY; PRT; 777 AA.
 AC O92RB3;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE PROBABLE GLUCOSE DEHYDROGENASE (PYROLOQUINOLINE-OUTINONE) PROTEIN
 DE (EC 1.1.99.17).

OS Rhizobium meliloti (Sinorhizobium meliloti).
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Rhizobiaceae; Sinorhizobium.
 NCBI_TaxID=382;
 RN
 RC SEQUENCE FROM N.A.
 RP STRAIN=1021;
 RX MEDLINE=21368234; PubMed=11474104;
 RA Gilbert F., Finan T.M., Long S.R., Puehler A., Abola P., Ampe F.,
 Barloy-Hubler F., Barnett M.J., Becker A., Bolstad P., Bothe G.,
 Boutry M., Bowser L., Bhurmesler J., Cadieu E., Capella D., Chai P.,
 Cowie A., Davis R.W., Dreno S., Federspiel N.A., Fisher R.F.,
 Gloux S., Godt J., Goffeau A., Golding B., Guzy J., Gurjal M.,
 Hernandez-Lucas I., Hong A., Hulzar L., Hymen R.W., Jones T., Kahn D.,
 Kahn M.L., Kaiman S., Keating D.H., Kiss E., Komp C., Lelaure Y.,
 Masny D., Palm C., Beck M.C., Pohl T.M., Portetelle D., Purnelle B.,
 Ramsperger U., Surzycki R., Thebaud P., Vandebol M.,
 Vorholter F.J., Weidner S., Wells D.H., Wong K., Yeh K.-C., Batut J.;
 RA "The composite genome of the legume symbiont Sinorhizobium meliloti";
 RT Science 293:668-672(2001).
 RL EMBL; AL591785; CAC4553.1;
 KW OXidoreductase; Complete proteome.
 SQ SEQUENCE 777 AA; 84174 MW; D2F478A3BB10430 CRC64;

Query Match 20.6%; Score 146.5; DB 16; Length 777;
 Best Local Similarity 38.5%; Pred. No. 7.5e-06;
 Matches 35; Conservative 9; Mismatches 34; Indels 13; Gaps 2;

35 NPPAGWISYGNQENYRHSPLTQITTEVNGOLQVLMARGMOPGKV-----QVT 83
 159 NVPPGEMHOGRTFRPGRTSPLDQIRPENVANLQVAMQ--YQGDVKLPEDEVSETTYQVT 216
 84 PLIHGVMYLANPGDVIAIDATGDLIWEHR 114
 217 PLKVKDTLVCTPHNMAIALDAATGQEKMKY 247

RESULT 11
 091115 PRELIMINARY; PRT; 803 AA.
 AC 091115;
 DT 01-MAR-2001 (TREMblrel. 16, Created)
 DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)
 DT 01-DEC-2001 (TREMblrel. 18, Last annotation update)
 DE GLUCOSE DEHYDROGENASE.
 GN GCD OR PA2290.
 OS Pseudomonas aeruginosa.
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 OC Pseudomonas.
 OX NCBI_TaxID=287;
 RN
 RC SEQUENCE FROM N.A.
 RP STRAIN=ATCC 15692 / PA01;
 RX MEDLINE=20437337; PubMed=10984043;
 RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,
 Hickey M.J., Brinkman F.S.L., Huftagle W.O., Kowalik D.J., Lagrou M.,
 Garber R.L., Goltzy L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
 Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
 Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
 Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
 RA "Complete genome sequence of Pseudomonas aeruginosa PA01, an
 opportunistic pathogen.";
 RT Nature 406:959-964(2000).
 RL EMBL; AE004654; AAG05678.1;
 DR InterPro; IPR002372; Bac_PQQ_repeat.
 DR Pfam; PF01011; Bacterial_PQQ; 7.
 KW Complete Proteome.
 SQ SEQUENCE 803 AA; 86217 MW; 7DDF46B0FE5ED249 CRC64;

Query Match 20.1%; Score 143; DB 16; Length 803;
 Best Local Similarity 26.9%; Pred. No. 1.7e-05;

Matches 47; Conservative 20; Mismatches 60; Indels 48; Gaps 6;
 6 LLM-----ASAGALLAAPAQYTPVDELANP----- 36
 103 LLMPLWFRRLPADGPAPLGALGAVVLAAGAAVSSQ-FTNPGQVIGRIDDSGMTST 161
 37 ----PAGEWISYGNQENYRHSPLTQITTEVNGOLQVLMARGMOPGKV-----Q 81
 162 AAPAPDDGQAVGTETFGDRSPKQITPANVGQLEBAMR--ITGDLPTPADPLELTNE 219
 82 VTPLIHGVMYLANPGDVIAIDATGDLIWEHRQPLNATINSFGEPT-RGMA 135
 220 NTPKVNGLVACTAHSKVALDLPDGTAEIWRPDQIQSPVGKGFAMTCRGVS 274

RESULT 12
 091115 PRELIMINARY; PRT; 601 AA.
 AC 091115;
 DT 01-MAR-2001 (TREMblrel. 16, Created)
 DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
 DE MEFHANOI DEHYDROGENASE LARGE SUBUNIT-LIKE PROTEIN.
 GN MAF.
 OS Rhizobium meliloti (Sinorhizobium meliloti).
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Rhizobiaceae; Sinorhizobium.
 OX NCBI_TaxID=382;
 RN
 RC SEQUENCE FROM N.A.
 RP STRAIN=RM1021;
 RA Fennel B.J., Tiwari R.P., Dilworth M.J.;
 RA "Regulation of CI assimilation in Sinorhizobium meliloti.";
 RT Submitted (SEP-2000) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AF309488; AAC31643.1;
 DR HSSP; P38539; 4AAH.
 DR InterPro; IPR002372; Bac_PQQ_repeat.
 DR Pfam; PF01011; Bacterial_PQQ; 7
 SQ SEQUENCE 601 AA; 65866 MW; 09F74372B5B2E2A83 CRC64;

Query Match 19.8%; Score 140.5; DB 2; Length 601;
 Best Local Similarity 31.8%; Pred. No. 2.1e-05;
 Matches 41; Conservative 29; Mismatches 50; Indels 9; Gaps 7;

14 LALLAAPAQYTPVDEL--LANPPAGWISYGNQENYRHSPLTQITTEVNGOLQV 71
 8 LAIMSIGGAQVAFANDELQKLIDP--NQMAIQGTVANLRSKLDQIKDKNVAKLQVAM 66
 72 --ARGMOPGKVQYTPLIHGVWYLANP-GDVIAID-ATGDLIWEHR-RQPLNATINS 126
 67 TETSTGVLRGH-EGSPVIGDLMVHTPPNYVALDLSKDGQIWKYEPKQDPNYVPVC 125
 127 FGEPTRGMA 135
 126 CDTVNRGVA 134

RESULT 13
 092WY9 PRELIMINARY; PRT; 601 AA.
 AC 092WY9;
 DT 01-DEC-2001 (TREMblrel. 19, Created)
 DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
 DE PUTATIVE MEFHANOI DEHYDROGENASE PROTEIN, LARGE SUBUNIT (EC 1.1.99.8).
 GN SMW20173.
 OS Rhizobium meliloti (Sinorhizobium meliloti).
 OG Plasmid pSymB (megaplasmid 2).
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Rhizobiaceae; Sinorhizobium.
 OX NCBI_TaxID=382;

RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-1021;
 RX MEDLINE=21396508; PubMed=11481431;
 RA Fidan T.M., Weidner S., Wong K., Buhrmester J., Chain P.,
 RA Vorhoelet F.J., Hernandez-Lucas I., Becker A., Cowie A., Gouzy J.,
 RA "The complete sequence of the 1,683-kb PSYMB megaplasmid from the N2-
 RT fixing endosymbiont *Sinorhizobium meliloti*."
 RL Proc. Natl. Acad. Sci. U.S.A. 98:9889-9894(2001).
 DR EMBL: AF603642; CAC48573.1;
 KW Oxidoreductase; plasmid; Hypothetical protein; Complete proteome.
 SQ SEQUENCE 601 AA; 65759 MW; D73424FEED15ADBE CRC64;

Query Match 19.8%; Score 140.5; DB 16; Length 601;
 Best Local Similarity 31.8%; Pred. No. 2.1e-05;
 Matches 41; Conservative 29; Mismatches 50; Indels 9; Gaps 7;

QY 14 LALIAAPFAQVTPVTDL--LANPAGEMISYGONENRHSPLQITTEVNGOLQVW 71
 DB 8 LAISIGGAQVAFANDELQKLIIDP--NQMAIQGVANLRYSKLDQINKDNVSKLOVAV 66
 QY 72 --ARGMOPGVQVTPPLIHGVMILANP-GVYQAIID-AKTGDIWEHR-RQLPNIATLNS 126
 DB 67 TTFSTGVLRGH-EGSPLVIGDLMVHTFPNTVYALDLSKDGQIMWKYEPKODENVIPVNC 125
 QY 127 FGEPTRGMA 135
 DB 126 CDTVNRGVA 134

RESULT 14
 Q9X2S5 PRELIMINARY; PRT; 790 AA.
 AC Q9X2S5;
 DT 01-NOV-1999 (TREMBLrel. 12, Created)
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE GLUCOSE DEHYDROGENASE.
 GN GDH.
 OS Pantoea citrea.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Pantoea.
 OX NCBI_TaxID=53336;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=1056R;
 RX MEDLINE=99303331; PubMed=10376838;
 RA Pujol C.J., Kado C.I.;
 RT "gdh, a gene encoding a second quinoprotein glucose dehydrogenase in
 RT Pantoea citrea, is required for pink disease of pineapple."
 RL Microbiology 145:1217-1226(1999).
 DR EMBL: AF050503; AAD23735.1;
 DR InterPro: IPR001479; Bac_P00.
 DR InterPro: IPR002372; Bac_P00_repeat.
 DR Pfam: PF01011; Bacterial_P00.
 DR PROSITE: PS00364; BACTERIAL_P00_2.1
 SQ SEQUENCE 790 AA; 85681 MW; 8BBB7625E0A5F089 CRC64;

Query Match 17.0%; Score 121; DB 2; Length 790;
 Best Local Similarity 32.6%; Pred. No. 0.0024;
 Matches 31; Conservative 13; Mismatches 39; Indels 12; Gaps 3;

QY 27 PYTDLANPAGEMISYGONENRHSPLQITTEVNGOLQVW-----RGMOPGK 79
 DB 152 PATQPVSAAGK--DWTAYGQTQAGDRYSSLSQINSNTKQNVAVMTYSGGVKRPDDGE 208
 QY 80 V--QVTPPLIHGVMILANP-GVYQAIIDAKTGDLIV 112
 DB 209 ITNESTPLKVGMLTYCTAHQILVAIDAATGKERK 243

RESULT 15
 ID P71509 PRELIMINARY; PRT; 601 AA.
 AC P71509;
 DT 01-FEB-1997 (TREMBLrel. 02, Created)
 DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE METHANOL DEHYDROGENASE LARGE SUBUNIT HOMOLOG.
 GN MXAP.
 OS Methylobacterium extorquens.
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Methylobacterium group; Methylobacterium.
 OX NCBI_TaxID=408;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=AM1;
 RX MEDLINE=97312011; PubMed=9168622;
 RA Chistoserdova L., Lidstrom M.E.;
 RT "Molecular and mutational analysis of a DNA region separating two
 RT methylotrophy gene clusters in *Methylobacterium extorquens* AM1."
 RL Microbiology 143:1729-1736(1997).
 DR EMBL: U72662; AAB58890.1;
 DR HSSP: P38539; AAH.
 DR InterPro: IPR002372; Bac_P00_repeat.
 DR Pfam: PF01011; Bacterial_P00.
 SQ SEQUENCE 601 AA; 64952 MW; 68E45C7059CBB239 CRC64;

Query Match 16.8%; Score 119; DB 2; Length 601;
 Best Local Similarity 29.3%; Pred. No. 0.0027;
 Matches 41; Conservative 29; Mismatches 58; Indels 12; Gaps 8;

QY 1 MKPTSLMASGALALAPFAQVTPVTDL--LANPAGEMISYGONENRHSPLQIT 60
 DB 1 MRAVHLALGAGLAA--ASPALANESVLRG--VANP--AEQVLOTVDYANTRYSKLDQIN 54
 QY 61 TENVGOLQVW--ARGMOPGVQVTPPLIHGVMILANP-GVYQAIIDAKTG-DLIWEHR- 115
 DB 55 ASNVKNLQVAVTFTSTGVLRGH-EGSPLVGNIMVHTFPNTVYALDLDGAKIWKYEP 113
 QY 116 RQLPNIATLNSFGEPTRGMA 135
 DB 114 KODPSVIPVNCDDTVNRGLA 133

Search completed: May 24, 2002, 10:17:32
 Job time: 745 sec

Fri May 24 11:27:20 2002

us-08-934-506a-5_copy_1_135.rspt

Page 7

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2002, 10:05:02 ; Search time 117.28 Seconds
(without alignments)
357.050 Million cell updates/sec

Title: US-08-934-506a-5_COPY_180_556
Perfect score: 2032
Sequence: 1 IVANGIVAVAGSTQGYSPFG.....GMQYVALAGGVSYSGSLNS 377

Scoring table:
BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 11073796 residues
Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
A.Geneseq_032802:*
1: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT:*
2: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT:*
3: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT:*
4: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT:*
5: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT:*
6: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT:*
7: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT:*
8: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT:*
9: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT:*
10: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT:*
11: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT:*
12: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT:*
13: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT:*
14: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT:*
15: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT:*
16: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT:*
17: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT:*
18: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT:*
19: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT:*
20: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT:*
21: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT:*
22: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2032	100.0	579	19	AAW37873 Alcohol and/or ald
2	1794	88.3	579	19	AAW37876 Alcohol and/or ald
3	1777	87.5	579	19	AAW37874 Alcohol and/or ald
4	1736	85.4	578	19	AAW37875 Alcohol and/or ald
5	338	16.6	742	11	AAW05235 Amino acid sequenc
6	332.5	16.4	738	13	AAW20192 ADH complex protei
7	331.5	16.3	738	11	AAW31993 A.aliceteligens me
8	326.5	16.1	754	21	AAW35987 Sorbitol dehydroge
9	202	9.9	740	20	AAW95019 Sorbitol dehydroge
10	108.5	5.3	530	22	AAW12236 Human PRO4340 poly
11	107	5.3	269	11	AAW07006 Protein G variant.

12	107	5.3	269	15	AAW53295 IgG-binding Strept
13	106.5	5.2	593	11	AAW07014 Protein G variant
14	106.5	5.2	594	12	AAW10005 Streptococcus GX78
15	106	5.2	269	10	AAW94788 Protein G variant.
16	106	5.2	715	22	AAW81738 S. epidermidis ope
17	105.5	5.2	429	22	AAW66542 Human interferon-a
18	105.5	5.2	429	22	AAW93799 Human protein sequ
19	105.5	5.2	469	21	AAW56578 Human protease can
20	105.5	5.2	278	8	AAW70468 Sequence of polype
21	105	5.1	480	22	AAW82625 S. epidermidis ope
22	103	5.0	549	21	AAW96642 B. pseudomallei Gr
23	102	5.0	790	22	AAW67230 Amino acid sequenc
24	99.5	4.9	391	21	AAW81628 Streptococcus pneu
25	99.5	4.9	770	22	AAW30824 Amino acid sequenc
26	99.5	4.9	910	22	AAW83007 S. epidermidis ope
27	99	4.9	696	22	AAW82914 S. epidermidis ope
28	99	4.9	886	22	AAW93996 Human protein sequ
29	98.5	4.8	228	21	AAW10433 Expression vector
30	98.5	4.8	593	15	AAW62944 Streptococcus Prot
31	98.5	4.8	1016	22	AAW38901 C. trachomatis CT8
32	98	4.8	7107	22	AAW58144 Drosophila melanog
33	97.5	4.8	524	20	AAW37238 Chlamydia trachoma
34	97.5	4.8	1315	20	AAW08642 S. aureus Strd pro
35	97.5	4.8	1349	22	AAW34402 Staphylococcus aur
36	97.5	4.8	1349	22	AAW37584 Fall armyworm Bt t
37	97.5	4.8	1734	22	AAW20500 Cosmid cHRIM5 enco
38	97	4.8	1673	21	AAW95707 Neisseria meningit
39	97	4.8	2599	22	AAW75098 Novel human diagno
40	96.5	4.7	396	22	AAW82859 Human tissue trans
41	96.5	4.7	548	21	AAW44918 S. epidermidis ope
42	96	4.7	470	22	AAW82701 Human tissue trans
43	95.5	4.7	548	15	AAW34013 C. trachomatis L2
44	95	4.7	505	20	AAW16739 Human GTP-binding
45	95	4.7	547	22	AAW68556

ALIGNMENTS

RESULT 1	AAW37873	AAW37873 standard; Protein: 579 AA.
ID	AAW37873	
AC	AAW37873	
XX		
DT	10-AUG-1998 (first entry)	
XX		
DE	Alcohol and/or aldehyde dehydrogenase A amino acid sequence.	
XX		
KW	Alcohol/aldehyde dehydrogenase A enzyme; recombinant organism;	
KW	aldehyde; ketone; carboxylic acid; L-sorbose; D-sorbitol;	
KW	2-keto-L-gulonic acid; L-ascorbic; inhibition.	
XX		
OS	Glucanobacter oxydans.	
XX		
FH	Key	Location/Qualifiers
FT	Peptide	1..23
FT		/note= "signal peptide"
FT	Protein	24..579
FT		/note= "mature protein"
XX		
PN	EP832974-A2.	
XX		
PD	01-APR-1998.	
XX		
PF	11-SEP-1997;	97EP-0115801.
XX		
PR	19-SEP-1996;	96EP-0115001.
XX		
PA	(HOFF) HOFFMANN LA ROCHE & CO AG F.	
XX		
PI	Asakura A, Hoshino T, Ojima S, Shinjoh M, Tomiyama N;	
XX		

DR WPI; 1998-195228/18.
 DR N-PSDB; AAV29051.
 PT Recombinant Gluconobacter oxydans alcohol and/or aldehyde
 PT dehydrogenase enzyme(s) - useful for converting substrate(s), e.g.
 PT L-sorbose or D-sorbitol to 2-keto-L-gulonic acid
 PS Claim 1; Pages 35-37; 59pp; English.
 XX This is the amino acid sequence for the Gluconobacter oxydans
 CC alcohol and/or aldehyde dehydrogenase A enzyme. The enzymes
 CC or recombinant organisms can be used to convert suitable substrates
 CC to aldehydes, ketones or carboxylic acids, especially to convert
 CC L-sorbose or D-sorbitol to 2-keto-L-gulonic acid, which can be
 CC converted to L-ascorbic acid by standard procedures. The derivatives
 CC of ADH enzymes have desired substrate specificity, higher affinity
 CC to a substrate, lower affinity to an inhibitory compound, higher
 CC stability against temperature and/or pH and higher catalytic speed.
 XX
 XX Sequence 579 AA:

Query Match 100.0%; Score 2032; DB 19; Length 579;
 Best Local Similarity 100.0%; Pred. No. 5.5e-178;
 Matches 377; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IVANGYIVAGSTCOYSPFGCFVSGHDSATGEELMRNYFIPRAGEEGDEMGNDYEARMWT 60
 DB 180 IVANGYIVAGSTCOYSPFGCFVSGHDSATGEELMRNYFIPRAGEEGDEMGNDYEARMWT 239
 QY 61 GAMQITTYDVTNLVHYGSTAVGPASERQGTGCTLYGTNTRFAVRPDTGEIWRHQT 120
 DB 240 GAMQITTYDVTNLVHYGSTAVGPASERQGTGCTLYGTNTRFAVRPDTGEIWRHQT 299
 QY 121 PRDMMDOECTFEEMVNTNVDOPSTMEGLQSTINPNAATGERVLTGVPCKTGMQFDAA 180
 DB 300 PRDMMDOECTFEEMVNTNVDOPSTMEGLQSTINPNAATGERVLTGVPCKTGMQFDAA 359
 QY 181 TGEFLMARDNTYONMIESIDENGIYVNEADAIKELDEYDVCPTFLGRDPSAALNPD 240
 DB 360 TGEFLMARDNTYONMIESIDENGIYVNEADAIKELDEYDVCPTFLGRDPSAALNPD 419
 QY 241 SGITFIPLNWVCYDMAVDQETSMQVNTSVTKLPKQKMGKIDAIIDISTGRTLSV 300
 DB 420 SGITFIPLNWVCYDMAVDQETSMQVNTSVTKLPKQKMGKIDAIIDISTGRTLSV 479
 QY 301 ERAAANYSPVLTSGGVLFGNGTDRYFRALSOETGETIMOTRLATVAGSOAISYEVDGM 360
 DB 480 ERAAANYSPVLTSGGVLFGNGTDRYFRALSOETGETIMOTRLATVAGSOAISYEVDGM 539
 QY 361 YVIAAGGVSYSGSLNS 377
 DB 540 YVIAAGGVSYSGSLNS 556

RESULT 2
 AAW37876
 ID AAW37876 standard; Protein: 579 AA.
 AC AAW37876;
 XX
 XX 10-AUG-1998 (first entry)
 DE Alcohol and/or aldehyde dehydrogenase B amino acid sequence.
 XX
 XX Alcohol/aldehyde dehydrogenase B enzyme; recombinant organism;
 KW aldehyde; ketone; carboxylic acid; L-sorbose; D-sorbitol;
 KW 2-keto-L-gulonic acid; L-ascorbic; inhibition.
 OS Gluconobacter oxydans.
 XX
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..23

FT /note= "signal peptide"
 FT 24..579
 FT Protein
 FT /note= "mature protein"
 PN EP832974-A2.
 XX
 XX 01-APR-1998.
 PD
 XX
 XX 11-SEP-1997; 97EP-0115801.
 PR 19-SEP-1996; 96EP-0115001.
 XX
 XX (HOFF) HOFFMANN LA ROCHE & CO AG F.
 PI Asakura A, Hoshino T, Ojima S, Shinjoh M, Tomiyama N;
 XX WPI; 1998-195228/18.
 DR N-PSDB; AAV29054.
 XX
 XX
 PT Recombinant Gluconobacter oxydans alcohol and/or aldehyde
 PT dehydrogenase enzyme(s) - useful for converting substrate(s), e.g.
 PT L-sorbose or D-sorbitol to 2-keto-L-gulonic acid
 PS Claim 1; Pages 44-46; 59pp; English.

This is the amino acid sequence for the Gluconobacter oxydans
 CC alcohol and/or aldehyde dehydrogenase B enzyme. The enzymes
 CC or recombinant organisms can be used to convert suitable substrates
 CC to aldehydes, ketones or carboxylic acids, especially to convert
 CC L-sorbose or D-sorbitol to 2-keto-L-gulonic acid, which can be
 CC converted to L-ascorbic acid by standard procedures. The derivatives
 CC of ADH enzymes have desired substrate specificity, higher affinity
 CC to a substrate, lower affinity to an inhibitory compound, higher
 CC stability against temperature and/or pH and higher catalytic speed.
 XX
 XX Sequence 579 AA:

Query Match 88.3%; Score 1794; DB 19; Length 579;
 Best Local Similarity 85.7%; Pred. No. 4.2e-156;
 Matches 323; Conservative 27; Mismatches 27; Indels 0; Gaps 0;

QY 1 IVANGYIVAGSTCOYSPFGCFVSGHDSATGEELMRNYFIPRAGEEGDEMGNDYEARMWT 60
 DB 181 IVANGYIVAGSTCOYSPFGCFVSGHDSATGEELMRNYFIPRAGEEGDEMGNDYEARMWT 240
 QY 61 GAMQITTYDVTNLVHYGSTAVGPASERQGTGCTLYGTNTRFAVRPDTGEIWRHQT 120
 DB 241 GAMQITTYDVTNLVHYGSTAVGPASERQGTGCTLYGTNTRFAVRPDTGEIWRHQT 299
 QY 121 PRDMMDOECTFEEMVNTNVDOPSTMEGLQSTINPNAATGERVLTGVPCKTGMQFDAA 180
 DB 300 PRDMMDOECTFEEMVNTNVDOPSTMEGLQSTINPNAATGERVLTGVPCKTGMQFDAA 359
 QY 181 TGEFLMARDNTYONMIESIDENGIYVNEADAIKELDEYDVCPTFLGRDPSAALNPD 240
 DB 360 TGEFLMARDNTYONMIESIDENGIYVNEADAIKELDEYDVCPTFLGRDPSAALNPD 419
 QY 241 SGITFIPLNWVCYDMAVDQETSMQVNTSVTKLPKQKMGKIDAIIDISTGRTLSV 300
 DB 420 SGITFIPLNWVCYDMAVDQETSMQVNTSVTKLPKQKMGKIDAIIDISTGRTLSV 479
 QY 301 ERAAANYSPVLTSGGVLFGNGTDRYFRALSOETGETIMOTRLATVAGSOAISYEVDGM 360
 DB 480 ERAAANYSPVLTSGGVLFGNGTDRYFRALSOETGETIMOTRLATVAGSOAISYEVDGM 539
 QY 361 YVIAAGGVSYSGSLNS 377
 DB 541 YVIAAGGVSYSGSLNS 556

RESULT 3
 AAW37874

```

ID AAM37874 standard; Protein; 579 AA.
AC AAM37874;
XX
XX 10-AUG-1998 (first entry)
DT
DE Alcohol and/or aldehyde dehydrogenase A' amino acid sequence.
DE
XX
XX Alcohol/aldehyde dehydrogenase A' enzyme; recombinant organism;
KM aldehyde; ketone; carboxylic acid; L-sorbose; D-sorbitol;
KM 2-keto-L-gulononic acid; L-ascorbic acid; inhibition.
XX
XX Gluconobacter oxydans.
OS
XX
XX Key Location/Qualifiers
FH Peptide 1..23
FT /note= "signal peptide"
FT 24..579
FT Protein /note= "mature protein"
XX
XX EP832974-A2.
XX
XX 01-APR-1998.
XX
XX 11-SEP-1997; 97EP-0115801.
XX
XX 19-SEP-1996; 96EP-0115001.
XX
XX (HOFF ) HOFFMANN LA ROCHE & CO AG F.
XX
XX Asakura A, Hoshino T, Ojima S, Shinjoh M, Tomiyama N;
PI WPI; 1998-195228/18.
XX N-PSDB; AAV29052.
DR
XX
XX Recombinant Gluconobacter oxydans alcohol and/or aldehyde
PT dehydrogenase enzyme(s) - useful for converting substrate(s), e.g.
PT L-sorbose or D-sorbitol to 2-keto-L-gulononic acid
XX
XX Claim 1; Pages 38-40; 59pp; English.
XX
XX This is the amino acid sequence for the Gluconobacter oxydans
CC alcohol and/or aldehyde dehydrogenase A' enzyme. The enzymes
CC or recombinant organisms can be used to convert suitable substrates
CC to aldehydes, ketones or carboxylic acids, especially to convert
CC L-sorbose or D-sorbitol to 2-keto-L-gulononic acid, which can be
CC converted to L-ascorbic acid by standard procedures. The derivatives
CC of ADH enzymes have desired substrate specificity, higher affinity
CC to a substrate, lower affinity to an inhibitory compound, higher
CC stability against temperature and/or pH and higher catalytic speed.
XX
XX Sequence 579 AA:
SQ

```

```

Query Match 87.5%; Score 1777; DB 19; Length 579;
Best Local Similarity 85.3%; Pred. No. 1.5e-154;
Matches 318; Conservative 29; Mismatches 26; Indels 0; Gaps 0;

```

```

QY 1 IVANGVIVAGSTCOYSPFCFVSGHDSATGGEELMRYFTIPRAGEEDGNDYEARMNT 60
DB 180 ivangvivaagstcypfcfvgshdsatggeelwryfiprgeedgndyearmnt 239
QY 61 GAWGQITVDYVNLVHYGSTAVGPASETORGPFGTLVGTNRFVAVRPDTGTVWRHQT 120
DB 240 gawgqitvdypvnlvhygstavgpasetorgpfgtlvgtnrfvavrpdgtvwrhqt 299
QY 121 PRONMDECFEFEMAVTNVQVSTEMEGLOSTINPNAATGERRVLTGVPCKTGMQFDE 180
DB 300 prdnmdectfefemavtnvqvstemeeglostinpnaatgeravlvgpcktgmmqfdae 359
QY 181 TGEFIARNTNYNMIESIDENGIIVTNEDAIKELDEYDVCPFTAGGRDWPSSALNDP 240
DB 360 tgefiarntnynmiesidengiivtnedaikeledeydvcpftaggrdwpssalndp 419

```

```

QY 241 SGVIFPLNNVCDMAVDQEFITSMDYNTSNVTKLPKDKMIGRIDAIDISTGRTIMSV 300
DB 420 tgiyifplnnvcdmavdqefitismdyntsnvtklpkdkmigridaidistgrtimsv 479
QY 301 ERANANSPVLTSGGVLFPNGDRVFRALSOEGETLNTGRLATVATVSGAISEVDVGMQ 360
DB 480 ereananspvltsggvlfpngdrvfralsoegetlntgtrlatvatvsgaisevdmq 539
QY 361 YVATAGGVSYS 373
DB 540 yvataggvsyts 552

```

```

RESULT 4
AAM37875
ID AAM37875 standard; Protein; 578 AA.
AC AAM37875;
XX
XX 10-AUG-1998 (first entry)
DT
DE Alcohol and/or aldehyde dehydrogenase A' amino acid sequence.
DE
XX
XX Alcohol/aldehyde dehydrogenase A' enzyme; recombinant organism;
KM aldehyde; ketone; carboxylic acid; L-sorbose; D-sorbitol;
KM 2-keto-L-gulononic acid; L-ascorbic acid; inhibition.
XX
XX Gluconobacter oxydans.
OS
XX
XX Key Location/Qualifiers
FH Peptide 1..23
FT /note= "signal peptide"
FT 24..578
FT Protein /note= "mature protein"
XX
XX
XX EP832974-A2.
XX
XX 01-APR-1998.
XX
XX 11-SEP-1997; 97EP-0115801.
XX
XX 19-SEP-1996; 96EP-0115001.
XX
XX (HOFF ) HOFFMANN LA ROCHE & CO AG F.
XX
XX Asakura A, Hoshino T, Ojima S, Shinjoh M, Tomiyama N;
PI WPI; 1998-195228/18.
XX N-PSDB; AAV29053.
DR
XX
XX Recombinant Gluconobacter oxydans alcohol and/or aldehyde
PT dehydrogenase enzyme(s) - useful for converting substrate(s), e.g.
PT L-sorbose or D-sorbitol to 2-keto-L-gulononic acid
XX
XX Claim 1; Pages 41-43; 59pp; English.
XX
XX This is the amino acid sequence for the Gluconobacter oxydans
CC alcohol and/or aldehyde dehydrogenase A' enzyme. The enzymes
CC or recombinant organisms can be used to convert suitable substrates
CC to aldehydes, ketones or carboxylic acids, especially to convert
CC L-sorbose or D-sorbitol to 2-keto-L-gulononic acid, which can be
CC converted to L-ascorbic acid by standard procedures. The derivatives
CC of ADH enzymes have desired substrate specificity, higher affinity
CC to a substrate, lower affinity to an inhibitory compound, higher
CC stability against temperature and/or pH and higher catalytic speed.
XX
XX Sequence 578 AA:
SQ

```

```

Query Match 85.4%; Score 1736; DB 19; Length 578;
Best Local Similarity 83.8%; Pred. No. 9e-151;
Matches 315; Conservative 28; Mismatches 33; Indels 0; Gaps 0;

```

```

QY 1 IVANGIVAGSTCOYSPFGCGVSGHDSATGEBELMRNYFIRAGEEGDETWGNDYEARMWT 60
Db 179 IVANGIVAGSTCOYSPFGCGVSGHDSATGEBELMRNYFIRAGEEGDETWGNDYEARMWT 238
QY 61 GAMQITTYDPVTNLVHGSTAVGPASSETORGTCGTLGNTNRPVAPDGTGELVWRHQT 120
Db 239 GVVGFITVDPVGGVTVHGSAVPAEIQGTGVGSMYGENTRFVAPGELVWRHQT 298
QY 121 PRDMWDECTFEMAVTVWDVOPSTEMEGLOSINPNMATGRRVLTGVPCGTGMWQFDE 180
Db 299 PRDMWDECTFEMAVTVWDVOPSTEMEGLOSINPNMATGRRVLTGVPCGTGMWQFDE 358
QY 181 TGEIARWDTNQNMTESIDENGIVTVNEDAILKEIDVEYDVCPTFLGGRWPSAALNDP 240
Db 359 TGEIARWDTNQNMTESIDENGIVTVNEDAILKEIDVEYDVCPTFLGGRWPSAALNDP 418
QY 241 SGYIFILNNVCYDMAVDOEFTSMQVYNTNNTKLPKCDMIGRIDAIDISGRRLMSV 300
Db 419 TGIYFIPLNACSGMTAVDQETSLDYNNVSLDKLSPSEMGRIDAIDISTGRLMSV 478
QY 301 ERAAANYSPVLTGGGVLFNGGTDRTYFRALSOETGELMOTRLATVAGSQAISYEVDGMO 360
Db 479 EYASNYAPVLTGGGVLFNGGTDRTYFRALSOETGELMOTRLATVAGSQAISYEVDGMO 538
QY 361 YVAIAGGVSYSGSLN 376
Db 539 YVAIAGGVSYSGSLN 554

RESULT 5
AAR05235
ID AAR05235 standard; protein; 742 AA.
XX
XX AAR05235;
XX
XX 04-AUG-1990 (first entry)
XX
XX Amino acid sequence of alcohol dehydrogenase (ADH) encoded by bases
DE 1-2229 of alcohol dehydrogenase (ADH) gene.
XX
XX Alcohol dehydrogenase (ADH) gene; Acetobacter pasteurianus IFO 3191;
KW Acetobacter aceti K1006 (FERM-7528); plasmid PAA721.
XX
XX Acetobacter aceti strain K1006 (FERM-7528).
OS
XX JP02000452-A.
XX
XX 05-JAN-1990.
XX
XX 30-OCT-1987; 87JP-0273190.
XX
XX 30-OCT-1987; 87JP-0273190.
XX
XX 30-OCT-1987; 87JP-0273190, JP-075069.
XX
XX (QPPP) QP CORP (QPPF-).
XX
XX WPI: 1990-047990/07.
XX
XX N-PSDB; AAQ91811.
XX
XX Cloning DNA, plasmid and microbe contg. it
XX contains alcohol dehydrogenase gene derived from Acetobacter
XX aceti K1006 strain
XX
XX Disclosure: Fig 4; 8pp; Japanese.
XX
XX Also new are a recombinant plasmid contg. its encoding DNA, and a
XX microorganism transformed with the plasmid.
XX A DNA fragment was obt. from A. aceti
XX strain K1006 and was transferred to an ADH-defective strain. The
XX resulting plasmid, PAA721, contg. its gene was inserted into A.
XX pasteurianus strain IFO 3191 by the triparent method. Its gene is
XX useful for improving Acetobacter culturing.

```

```

SQ Sequence 742 AA:
Query Match 16.6%; Score 338; DB 11; Length 742;
Best Local Similarity 26.2%; Pred. No. 4,4e-22;
Matches 114; Conservative 67; Mismatches 152; Indels 102; Gaps 18;

QY 2 VANGIVAGSTCOYSPFGCGVSGHDSATGEBELMRNYFIRAGEEGDETWGNDYEARMWT 47
Db 203 VAKGLVILGN--GGAEGFARGTSAFSAETGKIKWTFVTPNKNPHASADILMKA 260
QY 48 -ETWGNDEYARMWT-----GAMQITTYDPVTNLVHGSTAVGPASSETORGTCGTLV-- 98
Db 261 YKTWGP--KGAVVQGGGGLTWDLSIVYDPVSDILY---LAVG-----NGSPWNYKYRSE 309
QY 99 -GTN-----TRAVAPDGTGELVWRHQTLPKCDMIGRIDAIDISGRRLMSV 358
Db 310 GIGSLNLFISGIVALKPECEYVWHFQATPMQWQVTSVQIMLDPVK----- 358
QY 153 NFNAATGE--RRVLTGVPCGTGMWQFDETEGELMARDTNYQNMTESID-----EN 202
Db 359 -----GEMINIVVHAP--KNGFVYLDKTEGELISGKNYVQWANGIDPITGRPMYND 411
QY 203 GIVTVNEDAILKEIDVEYDVCPTFLGGRWPSAALNDPSGITYFIPLNVCY----- 253
Db 412 GYLTLNG-----KFWYGI--PSPILGAINFMAYSPKTHLYIPAHQIPFGYKNQYGF 463
QY 254 -----DMAVDOEFTSMQVYNTNNTKLPKCDMIGRIDAIDISGRRLMSVERAANYSP 309
Db 464 KPHASWNVGIDMCKNGIPDPE--ARTAYIKDLNGWLLAWDPVKMETVWKIDHKPVWNG 522
QY 310 VLTGGGVLFNGGTDRTYFRALSOETGELMOTRLATVAGSQAISYEVDGMOYVA----- 363
Db 523 ILATGGLIFGLANGELHAYDATNGSDLYKFDAGSLIAPPMYSVNGQYVAVERVWG 582
QY 364 ----IAGGVSYSGS 374
Db 583 GIYPSMGVYGRISG 597

RESULT 6
AAR20192
ID AAR20192 standard; protein; 738 AA.
XX
XX AAR20192;
XX
XX 16-APR-1992 (first entry)
XX
XX ADH complex protein (mol.wt. 72.000).
XX
XX Alcohol dehydrogenase; acetic acid; fermentation.
KW Acetobacter alioacetigenes NH-24.
XX
XX JP03266988-A.
XX
XX 27-NOV-1991.
XX
XX 26-MAR-1990; 90JP-0073440.
XX
XX 26-FEB-1990; 90JP-0042301.
XX
XX 26-MAR-1990; 90JP-0073440.
XX
XX (NAKA-) NAKANO SUTEN KK.
XX
XX WPI: 1992-019325/03.
XX
XX N-PSDB; AAQ20383.
XX
XX Alcohol dehydrogenase complex structural gene - used in plasmid
XX PT and enhancing efficiency of acetic acid fermentation for
XX PT transformed acetic acid bacteria
XX
XX Disclosure: Fig 3(1-3); 21pp; Japanese.

```

XX Acetobacter transformed with the sequence encoding this protein can
CC enhance the efficiency of acetic acid fermentation. The ADH complex
CC can be easily extracted from the bacteria and purified and it can be
CC used for the determination of an alcohol.
CC See also AAQ20383-84, and -86-88.
XX

SQ Sequence 738 AA;

Query Match	16.4%;	Score 332.5;	DB 13;	Length 738;
Best Local Similarity	26.8%;	Pred. No. 1.4e-21;		
Matches 108;	Conservative 53;	Mismatches 179;	Indels 63;	Gaps 10;

[illegible]

RESULT	7
AAR13993	
ID	AAR13993 standard; Protein; 738 AA.
AC	
XX	AAR13993;
XX	
DT	09-DEC-1991 (first entry)
XX	
XX	A.altoacetigenes membrane-bound ADH 72kD sub-unit.
XX	
XX	alcohol dehydrogenase complex; carboxylic acid production
XX	
OS	Acetobacter altoacetigenes.
XX	
PN	EP448969-A.
XX	
PD	02-OCT-1991.
XX	
PF	26-FEB-1991; 91EP-0102793.
XX	
PR	26-MAR-1990; 90JP-0073440.
PR	26-FEB-1990; 90JP-0042391.
XX	
XX	(NAKA-) NAKANO VINEGAR KK.
XX	
PI	Tamaki T, Takemura H, Tayama K, Fukaya M, Okumura H;
PI	Kawamura Y;
XX	
XX	WPI, 1991-289462/40.
DR	N-PSDB; AAQ13580.

XX Gene for membrane-bound alcohol dehydrogenase complex - obt'd.
 PR from *Acetobacter altoacetigenes*, used for prodn. of enzyme for
 PR converting alcohol to acid
 XX
 PS Disclosure; Fig 3; 36pp; English.

Total RNA was prepared from *Altoctocigenes* MH-24, digested with PstI and SmaI and ligated to PstI-SmaI cleaved pUC18. The ligation mixture was used to transform *E. coli* JM109. Probes were designed based on the N-terminal amino acid sequence of the ADH complex isolated from *A. altoctocigenes* (see AAQ13582-013584). The gene encoding the 72kD ADH subunit was isolated and sequenced. The directly sequenced N-terminal region of the purified 72kD sub-unit corresponds to the sequence beginning at residue 36 of the deduced sequence. This suggests that the first 35 N-terminal amino acids form a leader peptide involved in secretion of the mature 72kD protein. The deduced amino acid sequence has 77 per cent homology with the same enzyme from *A. acetii* K6033. See also AAQ13581.

Sequence	738 AA
50	

Query Match	16.3%;	Score 331.5;	DB 12;	Length 738;
Best Local Similarity	26.8%;	Pred. No. 1.7e-21;		
Matches 108; Conservative	52;	Mismatches 180;	Indels 63;	Caps 10;

```

QY 2 VANGVIVAGSNCQVSPFCC--FVSGHSAAGEELMRVYFIPRAGEED----- 47
Db 204 lakgrvllgn--ggseigargivsaafaetgkvdmffvypnpknepdaadsvlmmyk 261
QY 48 ETWAGMDIYEARWMTGA-----WQOITYDPVTLNVHYGSTAVGASFTQGRFGPGL 97
Db 262 qtwc-----ptgawtqggsggvtwasiydvpradvilygvqngspwnkyrseggkdn 314
QY 98 YCTNTRFAVRDPTGTETIWRHQTLPEDNMDOECTEDMAVTVNDVQSPTEMGLOSINPNA 157
Db 315 lfigstvalvklpetygvvwhfgetpmdqdfsfdsqjlmtdlpl----- 357
QY 158 TGEERVLNGVPCCKTQTMQOFDAETGEFMAADTVNQNINIESID-ENGIVTVNEDALIKEL 216
Db 358 ngetthvnharkngffryidantgetisgknyyyvnmwaasldpkrtprrpynpdalytl 417
QY 217 DVEYDVCPEFLIGRWPMSALNPDSGIYFELNNVCYDMNAVDQEFT-SMDVYNTS---N 272
Db 418 gkewyvipdpbglnmfaamaefpkrqglyvipdqypflytnqvvgfcpbdsowljldm 477
QY 273 VTKLPBG-----KDMIGRIDALDISTGRTLMSVEERAAANSPVLTSGGVLFPNGTGD 324
Db 478 kyvlpdspeakqafykdklkywlvawdpkqgaearvvdhikgpmwngilatqgdllfglan 537
QY 325 RYFRRLASQETGELIMOTRLATVAASGAQSAIEYEDGQVYAIAAG 367
Db 538 gefhaydatngsdllfhfaadsyilppvtyllangkygvaavevg 580

```

XX	RESULT	8
XX	AAB35987	
ID	AAB35987	standard; Protein; 754 AA.
XX		
XX	AAB35987;	
XX		
DT	01-MAR-2001	(first entry)
XX		
DE	Sorbitol dehydrogenase subunit 1	amino acid sequence.
XX		
XX	Sorbitol dehydrogenase; SDH; Gluconobacter oxydans; D-sorbitol,	
KW	L-sorbitose production; 2-keto-L-gulonate acid.	
XX		
OS	Gluconobacter oxydans.	
XX		
PN	MO200065066-A1.	
XX		

PD 02-NOV-2000.
 XX
 XX 23-APR-1999; 99WO-IB00736.
 XX
 XX 23-APR-1999; 99WO-IB00736.
 XX
 XX (CHOI/) CHOI E.
 XX (RHEE/) RHEE S.
 XX (LEE/) LEE E.
 XX
 XX Chai E, Rhee S, Lee E;
 XX
 XX MPI: 2000-687351/67.
 XX N-PSDB: AAC83153.
 XX
 XX Novel membrane-bound sorbitol dehydrogenase nucleic acid molecule
 XX isolated from gluconobacter suboxydans useful for the fermentative
 XX production of 2-keto-L-gulonic acid and L-sorbose from D-sorbitol
 XX
 XX Claim 1; Fig 8; 96pp; English.
 XX
 XX This invention relates to an isolated membrane-bound sorbitol
 XX dehydrogenase (SDH) from gluconobacter oxydans. The invention includes
 XX polynucleotides (AAC83153 - AAC83154) encoding the 3 subunits of SDH
 XX (AAB35987 - AAB35989). Also included in the invention are two
 XX polynucleotide sequences AAC83156 and AAC83157 which contain the subunit
 XX coding regions. The sorbitol dehydrogenase polynucleotide sequences are
 XX useful for producing L-sorbose from D-sorbitol and for increasing the
 XX production of 2-keto-L-gulonic acid by transforming a host cell,
 XX especially Gluconobacter with the DNA and selecting the transformed host
 XX cell. Note: An additional 66 DNA sequences (fragments of the SDH subunit
 XX genes) encoding fragments of SDH are specifically claimed, however these
 XX sequences are not given separately in the specification but are included
 XX in sequences AAC83156 and AAC83157.
 XX
 XX Sequence 754 AA;

Query Match 16.1%; Score 326.5; DB 21; Length 754;
 Best Local Similarity 25.3%; Pred. No. 5.1e-21;
 Matches 107; Conservative 64; Mismatches 182; Indels 67; Gaps 12;

QY 2 VANGVIVASTGTCQYSPFGC--FVSGHDSATGEBELMRNYFIRAGEEGD----- 47
 DB 199 jakgkvlign--ggaeifargfvtaydaetckmdwrtfvtvnpdkpdaasddvlnska 256
 QY 48 -ETWGNDEARMTG--AMGQITVDPVTNLVHGSTAVGPASETORGTGPGTLTGWTNR 103
 DB 257 yltwkggawkgqggggtvwdsllydvtldlylgvngspwnykftiseqgnlflgsl 316
 QY 104 FAVRPDGEIIVRHOTLPDRDWMDOECTFEMAVTVNDVOPSTMEGLDLSINPNAATGE-RR 162
 DB 317 valnpdctgkywhfgetpmdqdytsvqimalampv-----ngemth 359
 QY 163 VLTGVPCKTGTMOQDAETGFEFLWADRTNYOMIESIDE--NGIYVNEADAIKELDEYD 221
 DB 360 vlnhap--kngrfyllidatgkyfsgkytyenwnglndpvtgrpnypdaewltngrpwy 418
 QY 222 VCPFFLGKRMPSAALNDSGITFPLANNVCY-----DMKAVDOETSMVY 268
 DB 419 gipddlgghmaamaspqtklyvipaqvfydppqgfkahdswmlglldmknlgll 478
 QY 269 NTSN---VTKLPKGMIGRIDAIDISTGTLMSVERAANSPVSTGGVLFNGGTD 324
 DB 479 ddnbpqhakdaqlkdlkgwivawdpqkgaatlvdkhkpwmngllatagvllfgqlan 538
 QY 325 RYFPAISOETGELTMOQRLATVASGOAISYEVDGMOYVAI---AG-----GVSVSGS 374
 DB 539 gethayattgkdlftfpagatlapvtytangkqvavewvggyltfflgvartsg 598

RESULT 9
 AAW95019

ID AAW95019 standard; Protein; 740 AA.
 XX
 XX AAW95019;
 XX
 XX 21-MAY-1999 (first entry)
 XX
 XX Sorbitol dehydrogenase (SDH) protein sequence.
 XX
 XX Sorbitol dehydrogenase; SDH; open reading frame; ORF2; L-sorbose;
 XX D-sorbitol; vitamin C; enzyme.
 XX
 XX Gluconobacter suboxydans.
 XX
 XX Key Location/Qualifiers
 XX Peptide 1..24
 XX Protein /note="signal sequence"
 XX /note="mature protein"
 XX
 XX EP897984-A2.
 XX
 XX 24-FEB-1999.
 XX
 XX 13-AUG-1998; 98EP-0115231.
 XX
 XX 21-AUG-1997; 97EP-0114432.
 XX
 XX (HOFF) HOFFMANN LA ROCHE & CO AG F.
 XX
 XX Hosino T, Miyazaki T, Ojima S, Shinjoh M, Tomiyama N;
 XX MPI; 1999-134646/12.
 XX N-PSDB; AAX21501.
 XX
 XX New D-sorbitol dehydrogenase gene and recombinant protein - useful
 XX for production of L-sorbose, an intermediate in vitamin C production
 XX
 XX Claim 1; Fig 3A-D; 39pp; English.
 XX
 XX This represents a sorbitol dehydrogenase (SDH) protein. The DNA
 XX encoding the SDH enzyme also encodes an open reading frame (ORF2)
 XX product upstream of the SDH open reading frame, needed for SDH activity
 XX in vivo. Host cells transformed by a vector comprising the sorbitol
 XX dehydrogenase. Recombinant D-sorbitol dehydrogenase is useful for
 XX producing L-sorbose from D-sorbitol. L-sorbose is an important
 XX intermediate in vitamin C production.
 XX
 XX Sequence 740 AA;

Query Match 9.9%; Score 202; DB 20; Length 740;
 Best Local Similarity 21.2%; Pred. No. 1.4e-09;
 Matches 108; Conservative 51; Mismatches 144; Indels 206; Gaps 24;

QY 2 VANGVIVASTC-----QYSPFGCVSGHDSATGEBELMRNYFIRAGEEGDQNDYEA 56
 DB 273 vlnghvvnvheylldggrtwapsq-vlrgydaesgfvw-----awdvnsq 317
 QY 57 RMTGTAMGQITDPVTNLVHGSTAVGPASETORGTGPGTLTGWTNR-----FAY 106
 DB 318 r-----sqpay-rvvlaverrlpjldrtqg-----gsrlrdrnsaadyal 363
 QY 107 RPD-----TGEIIVRHOTLPDRDWMDOECTFEMAVTVNDVOPSTMEGLDLS 151
 DB 364 rsdaenksavvaldvktgspvfwfgchakdvwdydgatcl-----mdmpg--- 412
 QY 152 INPNAATGERRVLTGVPCKTGTMOQDAETGFEFLWADRTNYOMIESIDE---GIV- 205
 DB 413 --pdqgtvpallm--ptkrqgtfvldrrtgk-----pallpveerppapypvip 456
 QY 206 -----TYNEDAI-----LKEID----- 217


```

Db      457 gddrpsrptpwsvmnpallrtpolketudmgnmspdlqclcrlikfranyvgstfpvsvdkpw 516
Oy      218 VEYDVCPTFLGGRMPSPSALNPDSCGITYFIPLN-NCYCDMAAVDOEFSM----- 265
Db      517 Iey--pgynsgsdwsgmsydpqsgsliaawnltipmydqivtrckadsqglmpiddpnrk 573
Oy      266 -----DYNTSNVTKLPPGKDMIGRIPAIDISNG-RTLMS 299
Db      574 pggsgaegnsgamdtptyrlvtrfwdqy-egmmcnrpp-----YgmIadmkhkgkxvkw 628
Oy      300 -----VERAANTSPVLSTGGGVLENG-GTDRFRALSQETGTLW 339
Db      629 hplltararngwlpclplweligfpmngsvvtrvggllilgaatdqiraidentgkvww 668
Oy      340 QTRLATVASGOAISVEYDGMQVYIATAGG 368
Db      689 savlpbgsgaunpmtlyeanghbyvaimsgg 717

```

RESULT	10
ID	AAU12236
ID	AAU12236 standard; Protein; 530 AA.
XX	
AC	AAU12236;
XX	
DT	24-OCT-2001 (first entry)
XX	
DE	Human PRO4340 polypeptide sequence.
XX	
KW	Human secretory and transmembrane; PRO; mammalian; cancer; lung; breast; prostate; cervical; tumour necrosis factor-alpha; TNF-alpha; catalase; ear; proliferation; glucose; free fatty acid; skeletal muscle; adipocyte; A-peptide; factor VIIA; gene therapy.
KM	
XX	
OS	Homo sapiens.
XX	
PM	WO200140466-A2.
XX	
PD	07-JUN-2001.
XX	
PF	01-DEC-2000; 2000MO-US32678.
XX	
PR	01-DEC-1999; 99MO-US28301.
PR	01-DEC-1999; 99MO-US28634.
PR	02-DEC-1999; 99MO-US28551.
PR	02-DEC-1999; 99MO-US28564.
PR	02-DEC-1999; 99MO-US28565.
PR	09-DEC-1999; 99US-0170262.
PR	16-DEC-1999; 99MO-US30095.
PR	20-DEC-1999; 99MO-US30911.
PR	20-DEC-1999; 99MO-US30999.
PR	30-DEC-1999; 99MO-US31243.
PR	06-JAN-2000; 2000MO-US00277.
PR	06-JAN-2000; 2000MO-US00376.
PR	11-FEB-2000; 2000MO-US03565.
PR	18-FEB-2000; 2000MO-US04341.
PR	18-FEB-2000; 2000MO-US04342.
PR	22-FEB-2000; 2000MO-US04414.
PR	24-FEB-2000; 2000MO-US04914.
PR	24-FEB-2000; 2000MO-US05004.
PR	01-MAR-2000; 2000MO-US05601.
PR	20-MAR-2000; 2000MO-US07377.
PR	21-MAR-2000; 2000MO-US07532.
PR	30-MAR-2000; 2000MO-US08439.
PR	17-MAY-2000; 2000MO-US13705.
PR	22-MAY-2000; 2000MO-US14042.
PR	30-MAY-2000; 2000MO-US14941.
PR	02-JUN-2000; 2000MO-US15264.
PR	10-NOV-2000; 2000MO-US30873.
XX	
PA	(GERTH) GENENTECH INC.
XX	
PI	Baker KP, Beresini M, DeForge L, Desnoyers L, Filvaroff E, Gao W;

PI Gerilisen M., Goddard A., Godowski PJ, Gurney AL, Sherwood S;
 PI Smith V., Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;
 XX WPI: 2001-408281/43.
 DR N-PSDB; AAS21308.
 XX
 XX Isolated, secretory and transmembrane PRO polypeptide used to detect
 PI other PRO polypeptides, link bioactive molecules to cells expressing
 PI PRO polypeptides, and detect the presence of mammalian tumours e.g.
 PI lung, breast, prostate, cervical -
 XX
 PS Claim 12: Fig 130; 813pp; English.
 XX

AAU12172-AAU12446 represent novel human secretory and transmembrane PRO polypeptides. The PRO polypeptides are useful to detect other PRO polypeptides, to link bioactive molecules to cells expressing PRO polypeptides, to modulate biological activities of cells expressing PRO polypeptides, and to detect the presence of mammalian lung, colon, breast, prostate, rectal, cervical or liver tumours by comparing PRO polypeptide expression in a cell sample to that in a control sample. Some of the 275 sequences are also useful to stimulate the release of tumour necrosis factor-alpha (TNF-alpha) from human blood, the proliferation or differentiation of chondrocytes, the proliferation or gene expression in pericyte cells, the release of proteoglycans from cartilage, the proliferation of inner ear utricular supporting cells or of T-lymphocytes, the release of a cytokine from peripheral blood monocytes (PMBCs), or the proliferation of endothelial cells. Some of the PRO polypeptides may modulate glucose or free fatty acid uptake by skeletal muscle cells or by adipocytes; or inhibit binding of A-peptide to factor VIIa. The PRO polypeptides can be used in assays to identify molecules involved in binding interactions. The polynucleotides encoding PRO polypeptides can be used to generate probes, antisense RNA/DNA, transgenic or knock out animals and can be used in gene therapy.

530 AA
Sequence

Query Match	5.3%	Score 108.5	DB 22	Length 530
Best Local Similarity	24.3%	Pred. No. 0.33		
Matches 79; Conservative	38	Mismatches 109	Indels 99	Gaps 20

[illegible]

RESULT	11	.
AA07006		
ID	AA07006	standard; protein; 269 AA
XX		
AC	AA07006;	
XX		
DT	17-JAN-1991	(first entry)

[illegible]

Key	Location/Qualifiers
FT	Misc-difference 98 /note= "corresponds to AAA codon"
FT	Misc-difference 103 /note= "corresponds to GAA codon"
FT	Misc-difference 168 /note= "corresponds to GAA codon"
FT	Misc-difference 173 /note= "corresponds to GCT codon"
XX	US5312901-A.
XX	17-MAY-1994.
XX	14-FEB-1986; 86US-0829354.
XX	14-FEB-1986; 86US-0829354.
PR	14-FEB-1986; 86US-0829354.
PR	23-APR-1986; 86US-0854887.
PR	19-JUN-1987; 87US-0063959.
PR	20-JUN-1988; 88US-0209236.
PR	19-JUN-1990; 90US-0540169.
PR	21-APR-1992; 92US-0871539.
PA	(PHAA) PHARMACIA LKB BIOTECHNOLOGY AB.
XX	Fahnestock SR;
XX	WPI; 1994-159179/19.
DR	N-PSDB; AA064649.
XX	New recombinant streptococcal protein G variants - useful for antibody detection and purification and for therapy
XX	Claim 5; Column 47-48; 48pp; English.
XX	Protein G isolated from Streptococcus has IgG-binding activity which has been localised to the B repeating structure (see AAR35290).
XX	The sequence AAR35295 represents a claimed Streptococcal Protein G variant comprising the B domain active site and retaining IgG-binding activity.
XX	Sequence 269 AA;
XX	Query Match 5.3%; Score 107; DB 15; Length 269;
XX	Best Local Similarity 24.8%; Pred. No. 0.17;
XX	Matches 65; Conservative 29; Mismatches 110; Indels 58; Gaps 13
XX	1 IVANGVYAGSTCGQSPFCEVSHGDSATGELMRNYFIPRAGEBDETWGNDYEARWMT 60
XX	14 IINGKTLKGETT-----LEADAATAEKVFKY-----AND-----N 46
XX	61 GAMGGITVDPPVNLVHGSTAVGPASENGRNP-----GGTIVGNTNRFVAVRPDT 110
XX	47 GVDGGEWYLDADKTL--FTVTEKPEVIDASELPVAVTYKLVINGKTLKGETTLEAVDAAT 104
XX	111 GEIVWRHDTLPEDNM--DOCEPFEEMVTVTVQPTSTEMEGLOSINPNAATGERRVLTGVPC 169
XX	105 AEKVKFYA--NDNGVGDGEWYLDADKTLFTVTEKPEVIDASELP--AVTYKLVINGKTL 161
XX	170 KITGMWQ--FDAETGEFLWARTNTYONMIESIDENGI--VTVNEDAILKELDYDVCPTF 226
XX	162 KGETTKAVDAATAEKAFKYAN-----DNGVDGVWYLYDAS--TKTFTVTEMVTEVP 211
XX	LGGRWPSAALNPDGTYFIRPL 248

Db 212 rg--daptepekeasiplypl 231

RESULT 13

AA07014 standard; protein: 593 AA.

AA07014;
17-JAN-1991 (first entry)

Protein G variant with three active sites.

Immunoglobulin.

Streptococcus sp. Lancefield Group G strain.

Key Location/Qualifiers

Active-site 303..372

Active-site 373..427

Active-site 443..497

Active-site /label=B2

US4956296-A.

11-SEP-1990.

20-JUN-1988;

20-JUN-1988;

20-JUN-1988;

14-FEB-1986;

23-APR-1986;

17-FEB-1987;

19-JUN-1987;

(GENE-) GENEX CORP.

Fahnestock SR;

WPI: 1990-297491/39.

N-PSDB; AA006019.

Recombinant Protein G variants - obid. using a cloned gene

encoding Protein G from Streptococcus sp., used for binding

immunoglobulin.

Disclosure: Fig 9; 48pp; English.

Fragments and variants of the sequence are claimed esp. where

incorporated into a non-pathogenic host eg. E.coli, and expressed

at high levels.

The variants have a higher binding efficiency and capacity for

immunoglobulin, and may be used for purifying, detecting and

isolating antibodies.

Sequence 593 AA;

Query Match 5.2%; Score 106.5; DB 11; Length 593;

Best Local Similarity 25.4%; Pred. No. 0.6; Indels 55; Gaps 14;

Matches 66; Conservative 32; Mismatches 107; Indels 55; Gaps 14;

1 IVANGVIVAGSTCOYSPFGVSGHDSATGEELMKNYFIRPAGEGDETWGNDYFARMWT 60

306 IINGKTLKGETT-----leavdaataekvfky-----and-----n 338

61 GAWGQITTPDPVNTLVH-----YGSTAVGPASETORGTGPG-TLYGINTFRAPVPTDGE 112

339 gvdgwttydackttfvtkepevidaseltptvtykyingktlkgettleavdaatae 398

113 IYVRHQTLPFRDN-WDOECIFEMAVTVNDVQPSTEMEGLOSINPMNATGERRVLTGVPCKT 171

Db 399 kvfkgya--ndhvvdgewtyddatktftvtekepevidaselp-avtykyivngktlk 455

172 GTMMQ--FDAETGELWARPNTYONMIESIDENGI--VTVNEDAIKELDVEYDVPTEFG 228

456 ettikavdaetaekafkgyan-----dngvdgwvtydda-ekttvtamv--tevp 503

229 GRDWPMSALNPDGTYEIPDL 248

504 g-daptepekeasiplypl 522

RESULT 14

AA010005 standard; protein: 594 AA.

AA010005;

13-MAR-1991 (first entry)

Streptococcus GX7805 protein G.

Immunoglobulins; Ig.

Streptococcus sp GX7805.

Key Location/Qualifiers

Active-site 304..358

Active-site /label= Active Site B1

Active-site 374..428

Active-site 444..498

Active-site /label= Active Site B2

US4977247-A.

11-DEC-1990.

19-MAY-1989;

19-MAY-1989;

19-MAY-1989;

14-FEB-1986;

23-APR-1986;

17-FEB-1987;

19-JUN-1987;

20-JUN-1988;

(GENE-) GENEX CORP.

Fahnestock SR, Lee T, Wroble MH;

WPI: 1991-006758/01.

Q-PSDB; Q10002.

Immobilised protein G variants - used for detection, isolation

and purification. Immunoglobulin(s) and immunoglobulin fragments

Disclosure: Fig 9; 52pp; English.

Protein G gene product may be modified allowing the variant to

be immobilised and exhibit different binding profiles. The bound

protein is useful in purification and detection of Igs and fragments.

Sequence 594 AA;

Query Match 5.2%; Score 106.5; DB 12; Length 594;

Best Local Similarity 25.4%; Pred. No. 0.6; Indels 55; Gaps 14;

Matches 66; Conservative 32; Mismatches 107; Indels 55; Gaps 14;

1 IVANGVIVAGSTCOYSPFGVSGHDSATGEELMKNYFIRPAGEGDETWGNDYFARMWT 60

307 IINGKTLKGETT-----leavdaataekvfky-----and-----n 339

```

QY 61 GAMGQITVDPTNLNVH-----YGSNVAVGASSTQSGTPEG--LTVYGTNNRFRANRPTGE 112
Db 340 gvdgswtyddakktftvrekpevidasellpavtlytvpvngklkgeltteadaae 399
QY 113 IVMRHQTLPRDN--MDOECTFEPMATNVNDVOPSTMEGLQSIINPAATGERRVLTGVCKT 171
Db 400 kvfkqya--ndnvdgswtyddakktftvrekpevidasellp-avtlytlvngklkg 456
QY 172 GTMMQ--PDAETGEELMARDTYQNMIESIDENGI--VTVNEDATLKLDEYDVCPTEIG 228
Db 457 ettckavaaeataekafkqyan-----dnvgdywtlydda-ktfttcmv--teyp 504
QY 229 GRDMPAALNDPDSGITFPL 248
Db 505 g-dapltepekepaasipvlpl 523

```

[illegible]

Search completed: May 24, 2002, 10:05:04
Job time: 243 sec

XX	RESULT 15
XX	AAAP94788
XX	ID AAP94788 standard; protein; 269 AA.
XX	AC AAP94788;
XX	DT 04-JUL-1990 (first entry)
XX	DE Protein G variant.
XX	EW Protein G; Immunoglobulin; Fc receptor; ds.
XX	OS Streptococcus sp.
XX	PN W08810306-A.
XX	PD 29-DEC-1988.
XX	PF 20-JUN-1988; 88MO-US02084.
XX	PR 19-JUN-1987; 87US-0063959.
XX	PA (GENE-) GENEX CORP.
XX	PI Fahnstock SR;
XX	DR WPI: 1989-023848/03.
XX	DR N-PSDB; AAN94674.
XX	PT Cloned protein G variant genes -
XX	PT expressing proteins having immunoglobulin-binding properties of
XX	PT protein G and derived from Streptococcus sp.
XX	PS Claim 19; Page 89; 116pp; English.
XX	CC Gene for protein G variant of non-pathogenic streptococcus sp. allowing
XX	CC isolation of the protein and variants, useful as bacterial Fc receptors
XX	CC eg in purification and detection of Abs., screening of hybridoma clones
XX	CC and treatment of disease.
XX	Sequence 269 AA;

```

Query Match      5.2% Score 106; DB 10; Length 269;
Best Local Similarity 24.4%; Pred. No. 0.21;
Matches 64; Conservative 30; Mismatches 110; Indels 58; Gaps 13;

QY      1 IVANGVIVAGSTQYSPFCFVSGHDSATGEELANFYFIPRAGEGDETGNDYEARMW 60
      ::||::||::||::||::||::||::||::||::||::||::||::||::||
Db      14 lllngktllkyett-----teavdaaaekvixky-----and-----n 46

QY      61 GANGOIIVDPVPTNLVYGSTAVGAPASEQRGPP-----GGTLTGNTREFAVRPDT 110
      ||::||::||::||::||::||::||::||::||::||::||::||::||
Db      47 gvgdgewtyddakt--fvtlekpevidaseelpavtykllvngktllgettkvadaet 104

```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2002, 10:02:58 ; Search time 44.35 Seconds

(without alignments)
207.631 Million cell updates/sec

Title: US-08-934-506A-5_COPY_180_556
Perfect score: 2032
Sequence: 1 IVANGVIVAGSTGQSPFCG.....GMQYVAIAGGQVSYGSLNS 377

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: Issued_Patents_AA.*
2: /cgn2_6/ptodata/2/1aa/5A.COMB.pep.*
3: /cgn2_6/ptodata/2/1aa/5B.COMB.pep.*
4: /cgn2_6/ptodata/2/1aa/6A.COMB.pep.*
5: /cgn2_6/ptodata/2/1aa/6B.COMB.pep.*
6: /cgn2_6/ptodata/2/1aa/Backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	332.5	16.4	738	1	US-07-985-458-3
2	326.5	16.1	720	4	US-09-296-284-25
3	326.5	16.1	754	4	US-09-296-284-4
4	207	10.2	739	3	US-09-136-251-2
5	105	5.2	1012	1	US-08-219-262B-10
6	105	5.2	1012	3	US-09-031-655-10
7	95.5	4.7	548	5	PCT-US93-10541-2
8	95	4.7	1012	1	US-08-216-276A-19
9	94.5	4.7	512	1	US-09-356-818A-2
10	93.5	4.6	548	1	US-08-247-902A-2
11	93	4.6	380	3	US-08-971-782-4
12	93	4.6	380	4	US-09-309-026-4
13	93	4.6	459	3	US-08-971-782-2
14	93	4.6	459	4	US-09-309-026-2
15	91.5	4.5	606	2	US-08-883-534-3
16	91.5	4.5	606	3	US-09-204-764-3
17	91.5	4.5	816	1	US-07-731-157A-4
18	91.5	4.5	816	2	US-08-229-444B-2
19	91.5	4.5	816	2	US-08-541-780-4
20	91.5	4.5	1612	1	US-08-169-927-2
21	91	4.5	551	2	US-08-793-229-32
22	91	4.5	551	3	US-09-285-957-32
23	90.5	4.5	322	1	US-08-216-276A-29
24	90.5	4.5	687	5	PCT-US91-09784-2
25	90	4.4	484	1	US-08-216-276A-17
26	90	4.4	1012	1	US-08-219-262B-1
27	90	4.4	1012	1	US-08-219-262B-12

28	90	4.4	1012	3	US-09-031-655-1	Sequence 1, Appl
29	90	4.4	1012	3	US-09-031-655-12	Sequence 12, Appl
30	90	4.4	1222	2	US-08-682-517-15	Sequence 9, Appl
31	90	4.4	1252	2	US-08-682-517-9	Sequence 2, Appl
32	88	4.3	1012	1	US-07-944-943-2	Sequence 2, Appl
33	88	4.3	1012	1	US-08-219-262B-2	Sequence 2, Appl
34	88	4.3	1012	3	US-09-031-655-2	Sequence 2, Appl
35	88	4.3	1012	3	US-08-200-232-2	Sequence 2, Appl
36	88	4.3	1287	5	PCT-US95-02219A-2	Sequence 2, Appl
37	88	4.3	1287	5	PCT-US95-02219A-2	Sequence 2, Appl
38	87.5	4.3	592	1	US-08-217-327-8	Sequence 8, Appl
39	87.5	4.3	635	4	US-08-931-608A-5	Sequence 13, Appl
40	87	4.3	395	4	US-08-856-841-13	Sequence 22, Appl
41	87	4.3	395	4	US-08-856-841-22	Sequence 3, Appl
42	85.5	4.2	723	1	US-08-290-937B-3	Sequence 8, Appl
43	85	4.2	774	4	US-09-346-237-8	Sequence 2, Appl
44	85	4.2	5215	4	US-09-105-537-2	Sequence 2, Appl
45	84	4.1	500	2	US-08-987-519-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-07-985-458-3
Sequence 3, Application US/07985458
Patent No. 5344777
GENERAL INFORMATION:
APPLICANT: Tamaki, Toshimi;
APPLICANT: Takemura, Hiroshi;
APPLICANT: Fukaya, Kenji;
APPLICANT: Okumura, Hajime and
APPLICANT: Kawamura, Yoshiya
TITLE OF INVENTION: Structural Gene of Membrane-Bound
TITLE OF INVENTION: Alcohol Dehydrogenase Complex, Plasmid
TITLE OF INVENTION: Containing The Same And Transformed Acetic Acid
NUMBER OF INVENTION: 4
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fishauf, Holtz, Goodman & Woodward, P.C.
STREET: 600 Third Avenue
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10016-2088
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 0.72 mb
COMPUTER: IBM PC compatible (NEC PC-9801 ES)
OPERATING SYSTEM: MS DOS
SOFTWARE: ASCII Form
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/985,458
FILING DATE: 19921203
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/658,221
FILING DATE: 20-FEB-1991
APPLICATION NUMBER: 73440/1990
FILING DATE: 26-MAR-1990
ATTORNEY/AGENT INFORMATION:
NAME: Goodman, Herbert
REGISTRATION NUMBER: 17081
REFERENCE/DOCKET NUMBER: 910134/HG
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)972-1400
TELEFAX: (212)370-1622
TELEX: 236268
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 738 amino acids
TYPE: amino acid
TOPOLOGY: unknown

MOLECULE TYPE: protein
FEATURE:
NAME/KEY: MATURE PEPTIDE
LOCATION: 36 to 738
IDENTIFICATION METHOD: N-terminal sequences of the
IDENTIFICATION METHOD: purified protein having a molecular weight of about
IDENTIFICATION METHOD: 72,000
ORGANISM: Acetobacter alioacetigenes
STRAIN: MH-24
PUBLICATION INFORMATION:
AUTHORS: Tamaki, Toshimi;
AUTHORS: Fukaya, Masahiro;
AUTHORS: Takemura, Hiroshi;
AUTHORS: Tayama, Kenji;
AUTHORS: Okumura, Hajime;
AUTHORS: Kawamura, Yoshiya;
AUTHORS: Nishiyama, Makoto;
AUTHORS: Horinouchi, Sueharu and
AUTHORS: Beppu, Teruhiko
TITLE: Cloning and Sequencing of the Gene Cluster
TITLE: Encoding Two Subunits of Membrane-Bound
TITLE: Alcohol Dehydrogenase from Acetobacter
JOURNAL: Biochimica et Biophysica Acta.
VOLUME: 1088
PAGES: 292-300
DATE: 1991
US-07-985-458-3

Query Match 16.4%; Score 332.5; DB 1; Length 738;
Best Local Similarity 26.8%; Pred. No. 5.7e-25;
Matches 108; Conservative 53; Mismatches 179; Indels 63; Gaps 10;

QY 2 VANGVIAGSTCOYSPFGC--FVSGHDSATGEELMRNFIPRAGEGD----- 47
DB 204 IAKGVIIIGN--GGSEFGARGFVSAPDAETGKVDWRFETPRPKNEPDASDVLNKA 261
QY 48 ETWGNDEYEAHMTG-----WGQITDPVTNLVHYGSTAVGPASETGRTGGTL 97
DB 262 QTWS-----PTGAMTRGGGGGTWDSIVYDPADLVYLGVNGSPMNYKYRSEKGD 314
QY 98 YGTFEAVRPDGEIWMHQTLPDNDQECTEFEMMTNVNVOPTSTEMGLQSTINPMA 157
DB 315 LPLGSLVALAKPETGEYVHFQETPMQWDETSDDQIMTLDEPT----- 357
QY 158 TGERVYLGVPCKGTMTMOFDAETGEFLMARDTNTQNMIESID-ENGIVTVNEDAIL 216
DB 358 NGETRHVYIYHARKNGEFYIDAKTGEFISKNYVYVNNASGLDPRGTINPDLTYLT 417
QY 217 DVEYVCPFTLGRDMPKPSALNPDSCGTFIPLNVCYMMMAVDOEFT-SMDVYMTS--N 272
DB 418 GKEMWGIKPDGLGHNFAAMAFSPKGLVYIPAQGVPEFLYTVQVGGFTPHDPSMNLG 477
QY 273 VKLPPG-----KDMIGRIDIDISTGRITMSVEEAAANYSPVLSTGGCVLFNGTD 324
DB 478 KVGIDSEAKQAFYKDKLGMIVAMPDPOKQAEAMRVDDHKGPWNGGILATYGGDLF 537
QY 325 RYFRALSGEETLMQRLATVAGSGLSYEVDSQYVAIAG 367
DB 538 GEFHAYDATNGSDLFHFAADSGIIPVYTYLANGKQYVAIVEG 580

RESULT 2
US-09-296-284-25
; Sequence 25, Application US/09296284A
; Patent No. 6204040
; GENERAL INFORMATION:
; APPLICANT: Choi, Eun-Sung
; APPLICANT: Rhee, Sang-Ki
; APPLICANT: Lee, Eun-Hae
; TITLE OF INVENTION: Gluconobacter suboxydans sorbitol dehydrogenase, Genes

TITLE OF INVENTION: and Methods of Use Thereof
FILE REFERENCE: 1533.0870000
CURRENT APPLICATION NUMBER: US/09/296,284A
CURRENT FILING DATE: 1999-04-22
NUMBER OF SEQ ID NOS: 87
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 25
LENGTH: 720
TYPE: PRT
ORGANISM: Gluconobacter suboxydans
US-09-296-284-25

Query Match 16.1%; Score 326.5; DB 4; Length 720;
Best Local Similarity 25.5%; Pred. No. 2.2e-24;
Matches 107; Conservative 64; Mismatches 182; Indels 67; Gaps 12;

QY 2 VANGVIAGSTCOYSPFGC--FVSGHDSATGEELMRNFIPRAGEGD----- 47
DB 165 IAKGVIIIGN--GGSEFGARGFVSAPDAETGKVDWRFETPRPKNEPDASDVLNKA 222
QY 48 ETWGNDEYEAHMTG-----WGQITDPVTNLVHYGSTAVGPASETGRTGGTLGTR 103
DB 223 YPTMGKSGAMKQGGGGTWDLSLYDPTDLYLVYLGVNGSPMNYKYRSEKGNLFLGSI 282
QY 104 FAVRPDGEIWMHQTLPDNDQECTEFEMMTNVNVOPTSTEMGLQSTINPMA 162
DB 283 VAINDDTGKVMYHFEETPMQWDETSYQIIMALDMPV-----NGEMRH 325
QY 163 VLTGVPCKGTMTMOFDAETGEFLMARDTNTQNMIESIDE-NGIVTVNEDAILKELDEYD 221
DB 326 VLVNAP-KNGEYIIDAKTKFKISGKRYTENMANGLDPTGRPNYDPLMTLNGKPMY 384
QY 222 VCPFTLGRDMPKPSALNPDSCGTFIPLNVCY-----DMMAVDOEFTSMY 268
DB 385 GIPDGLGHNFAAMAFSPKGLVYIPAQGVPEFLYTVQVGGFTPHDPSMNLG 444
QY 269 NTSN-----VTKLPPGKMGICRIDIDISTGRITMSVEEAAANYSPVLSTGGCVLF 324
DB 445 DDNDPOHKAADKAPFLKDKLGMIVAMPDPOKQAAFTVDHKKPWWNGGLATYGGDLF 504
QY 325 RYFRALSGEETLMQRLATVAGSGLSYEVDSQYVAIAG-----AG-----GCVSYGSG 374
DB 505 GEFHAYDATNGSDLFHFAADSGIIPVYTYLANGKQYVAIVEG 564

RESULT 3
US-09-296-284-4
; Sequence 4, Application US/09296284A
; Patent No. 6204040
; GENERAL INFORMATION:
; APPLICANT: Choi, Eun-Sung
; APPLICANT: Rhee, Sang-Ki
; APPLICANT: Lee, Eun-Hae
; TITLE OF INVENTION: Gluconobacter suboxydans sorbitol dehydrogenase, Genes
; FILE REFERENCE: 1533.0870000
; CURRENT APPLICATION NUMBER: US/09/296,284A
; CURRENT FILING DATE: 1999-04-22
; NUMBER OF SEQ ID NOS: 87
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 4
; LENGTH: 754
; TYPE: PRT
; ORGANISM: Gluconobacter suboxydans
US-09-296-284-4

Query Match 16.1%; Score 326.5; DB 4; Length 754;
Best Local Similarity 25.5%; Pred. No. 2.4e-24;
Matches 107; Conservative 64; Mismatches 182; Indels 67; Gaps 12;

QY 2 VANGVIAGSTCOYSPFGC--FVSGHDSATGEELMRNFIPRAGEGD----- 47


```

Db 199 TAKGVIITGN--GGAEFGARGFVTAIDETGKMDKREFTTVNPNPKPGGAASDDVLSKA 256
QY 48 -ETGNDYEAARMGTG---AMGQITVPVTLNVHGSTAVGPASETGCTPGCTIGTNTFR 103
Db 257 YETWCKGAMKQGGGGGTVMDSLYDPTDLVYLGVNGSPWNYEFRESEGNLFLGSI 316
QY 104 FAVRPDTEIYWRHQTLPDRDNDQECTEMAVTNDVQPSSTEMELOSINNAATGE-RR 162
Db 317 VAINPDGTGYVWFQETPMDQMDYTSVOQIMALDMPV-----NGENRH 359
QY 163 VLTGVPCKTGMQDFAETGEFLMARDNTYQNMIESIDE-NGIYVNDAILKEIDVEYD 221
Db 360 VLYVHAP-KNGEFTYIIDAKTGKISKPTTYENWANGLDPRYGRPNYNDALMTLNGKPMY 418
QY 222 VCPTEGGRDMSALNPDGSIYFIPLNNVCY-----DMAVDOETSMQY 268
Db 419 GJPGDLGHNFAAMAYSPQTKLYIIPAOQVPFVDPKGGFKAHHDWNGLGDMNKICIL 478
QY 269 NTSN-----VTKLPGKDMIGRIDALDISTGRTLMSEVERAANYSPVLSTGGVLFNGGTD 324
Db 479 DUNDQHKRAKQFLKDLKGWIVAMPDQKQAFVVDHKGPMNGGLLTAQGVLFQGLAN 538
QY 325 RYFRALSOETGELMOTRLATVASGQALSYEVDGMQYVAI---AG-----GGVSYSG 374
Db 539 GFHNAVDTATGKDLFTFPQASLIAPVPTYTANGKQYVAVEVGMGCIYVFLGVAARTSG 598

```

RESULT 4
US-09-136-251-2
Sequence 2, Application US/09136251A

```

: Patent No. 6127156
: GENERAL INFORMATION:
: APPLICANT: HOSHINO, Tatsuo
: APPLICANT: MIYAZAKI, Taro
: APPLICANT: OJIMA, Setsuko
: APPLICANT: SHINJOH, Masako
: APPLICANT: TOMIYAMA, No. 6127156b1um1
: TITLE OF INVENTION: D-Sorbitol dehydrogenase gene
: FILE REFERENCE: D-Sorbitol dehydrogenase gene
: CURRENT APPLICATION NUMBER: US/09/136,251A
: EARLIER APPLICATION NUMBER: EP 97114432.4
: EARLIER FILING DATE: 1997-08-21
: NUMBER OF SEQ ID NOS: 9
: SOFTWARE: Patentin Ver. 2.1
: SEQ ID NO 2
: LENGTH: 739
: TYPE: PRT
: ORGANISM: Gluconobacter suboxydans
: FEATURE:
: NAME/KEY: SIGNAL
: LOCATION: (1)..(24)
: US-09-136-251-2

```

Query Match 10.2%; Score 207; DB 3; Length 739;
Best local similarity 21.5%; Pred. No. 3, le-12;
Matches 109; Conservative 51; Mismatches 142; Indels 206; Gaps 24;

```

QY 2 VANGVIYAGSTC-----QISPFCEVSGHDSATGEBELMRNYFIPRAGEGDEDTWGDYEA 56
Db 273 VINGVVVNEHVLIDGGRMAPSG-VIRGYDAESGKFEV-----AMDVNNSG 317
QY 57 RMMTGAGQITVDPVTLVHGSTAVGPASETGCTPGCTIGTNTFR-----FAV 106
Db 318 R-----SQPAY-RVYITVAVERRIGLPRRORG-----GSRLRPRDMSAADYLSL 363
QY 107 RPD-----TGEIYWRHQTLPDRDNDQECTEMAVTNDVQPSSTEMELOS 151
Db 364 RSDAENKVSAAVVAIDVKTGSPRWFQTAHRDWDYDIGSOATL-----MDMG--- 412
QY 152 INPMAATGERRVLTGVCKTGTMQFDATGELFMARDNTYQNMIESIDEN-----GIY- 205

```

```

Db 413 --PDGQIVPALIM---PTKRGQTFVLDRTGK-----PILPVEERAPSPGVI 456
QY 206 -----TVNDAI-----LIKED----- 217
Db 457 GDBPSPTQPSVGMFALRYVDLKETDWMGMSPIDQFCRIKFRANVYGFPTPSVDKPR 516
QY 218 VEYDVCPTEFLGGRDMSALNPDGSIYFIPLN-NVCYDMAVDOETSM----- 265
Db 517 IEY---PGYNGSGPWSMSVSDPQSGILIANMNITPMYDQLVTRKADSLGIMPIDPNFK 573
QY 266 -----DYNTSNVTKLPGKDMIGRIDALDISTG-RTIWS 299
Db 574 PGCGAGANGAMDTGPIYIVTPEFMDY-TGMKCNRP-----YGMTAIDMKHGGQVLMQ 628
QY 300 -----VERAANYSPVLSTGGGVLFNG-GTDYFRALSOETGELM 339
Db 629 HPLGTARANGFWGLPGLPWFELGTPNNGSGVYVGGGIFIGATDNGIRAIIDHTGKYVW 688
QY 340 QTRLATVASGQALSYEVDGMQYVAIAG 367
Db 689 SAVLPGGQANPMTYEANGHOYVAIAG 716

```

RESULT 5
US-08-219-262B-10
Sequence 10, Application US/08219262B

```

: Patent No. 5788970
: GENERAL INFORMATION:
: APPLICANT: VAKHARIA, VIKRAM
: APPLICANT: SNIDER, DAVID B
: APPLICANT: MENDEL-WHERSAT, STEPHANIE A
: TITLE OF INVENTION: CHIMERIC INFECTIOUS BURSAL DISEASE VIRUS
: TITLE OF INVENTION: CDNA CLONES, EXPRESSION PRODUCTS AND VACCINES BASED
: TITLE OF INVENTION: THEREON
: NUMBER OF SEQUENCES: 15
: CORRESPONDENCE ADDRESS:
: ADDRESS: OBION, SPIYAK, MCCLELLAND, MAIER & NEUSTADT
: STREET: 1755 S. JEFFERSON DAVID HIGHWAY, FOURTH FLOOR
: CITY: ARLINGTON
: STATE: VIRGINIA
: COUNTRY: USA
: ZIP: 22202
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/219,262B
: FILING DATE: 29-MAR-1994
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: OBION, NORMAN F
: REGISTRATION NUMBER: 24,618
: REFERENCE/DOCKET NUMBER: 2747-047-27
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (703) 413-3000
: TELEFAX: (703) 413-2220
: TELEX: 248855 OPAT UR
: INFORMATION FOR SEQ ID NO: 10:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1012 amino acids
: TYPE: amino acid
: STRANDEDNESS:
: TOPOLOGY: unknown
: MOLECULE TYPE: protein
: ORIGINAL SOURCE:
: ORGANISM: Infectious bursal disease virus
: STRAIN: OH
: US-08-219-262B-10

```

Query Match 5.2%; Score 105; DB 1; Length 1012;
 Best Local Similarity 22.2%; Pred. No. 0.12;
 Matches 84; Conservative 45; Mismatches 130; Indels 120; Gaps 19;

QY 71 VTNLVHGSTAV-----GPAS-----ETGRTPGTLYGNTNF 104
 DB 1 MTNIMDHTQOIVPFRSLMPTGTPASIPDDTLEKHTLRSESTYMLTVGDGSGILVFF 60
 QY 105 AVRPDGEIWMRHQTLPRD---NMDECTFEMMT-----NDVQPSFE 145
 DB 61 PGFP--GSVGAHYTLQSGSYQFDQ---MLTQNLPSVSYNCRVSRSLTVASSTL 113
 QY 146 MEGLOSINP--NAATGERRVLTGVPCKTGMQFDAETGEFLMARDTNYNMIESI---- 199
 DB 114 PGVYALNGTINAVT-----FGSLSEL---TDYSNGLSKATANIN 152
 QY 200 DENGIVYVNEAIIKEIDVEYDCPTFLGGRDWPASALNDPSGIYFIPLNNVC----- 252
 DB 153 DKIGNVIVGEGVTVLSPTSYDLSYRLGD-PIPAAGLDP-----KLMTCDSSDRPR 204
 QY 253 -YDMAAVDQ-EF-----TSMVYNTSNVTKLPPGKDM-GRIDAIDISTGRTLM- 298
 DB 205 VYTVTADEYQSSQLIPSGVKTTLFTANIDALTSVGGELIFSGVTHSIEVDVITYF 264
 QY 299 -----SVERAANYSPVLSGGLVFN--GCTDRYFRALSOETGETLMOTRLATVAS 348
 DB 265 IGFDTGEVTVKAVATDFGLTGTNNLVLPNLGGPTSEITQPTSMKLEVVYTKRGGT--A 322
 QY 349 GOAISYEVDGMQYVAIAGG 367
 DB 323 GDPISWTVSGTLAVTIYVG 341

RESULT 6
 US-09-031-655-10
 ; Sequence 10, Application US/09031655
 ; Patent No. 6017759
 ; GENERAL INFORMATION:
 ; APPLICANT: VAKHARIA, VIKRAM
 ; APPLICANT: SNYDER, DAVID B
 ; APPLICANT: MENDEL-WHERSAT, STEPHANIE A
 ; TITLE OF INVENTION: CHIMERIC INFECTIOUS BURSAI DISEASE VIRUS
 ; TITLE OF INVENTION: CDNA CLONES, EXPRESSION PRODUCTS AND VACCINES BASED
 ; TITLE OF INVENTION: THEREON
 ; NUMBER OF SEQUENCES: 15
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT
 ; STREET: 1755 S. JEFFERSON DAVID HIGHWAY, FOURTH FLOOR
 ; CITY: ARLINGTON
 ; STATE: VIRGINIA
 ; COUNTRY: USA
 ; ZIP: 22202
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/031,655
 ; FILING DATE:
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/219,262
 ; FILING DATE: 29-MAR-1994
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: OBLON, NORMAN F
 ; REGISTRATION NUMBER: 24,618
 ; REFERENCE/DOCKET NUMBER: 2747-047-27
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (703) 413-3000
 ; TELEFAX: (703) 413-2220
 ; TELEX: 248855 OPAT UR
 ; INFORMATION FOR SEQ ID NO: 10:

SEQUENCE CHARACTERISTICS:
 LENGTH: 1012 amino acids
 TYPE: amino acid
 STRANDEDNESS:
 TOPOLOGY: unknown
 MOLECULE TYPE: protein
 ORIGINAL SOURCE:
 ORGANISM: Infectious bursal disease virus
 SPRAIN: OH
 US-09-031-655-10

Query Match 5.2%; Score 105; DB 3; Length 1012;
 Best Local Similarity 22.2%; Pred. No. 0.12;
 Matches 84; Conservative 45; Mismatches 130; Indels 120; Gaps 19;

QY 71 VTNLVHGSTAV-----GPAS-----ETGRTPGTLYGNTNF 104
 DB 1 MTNIMDHTQOIVPFRSLMPTGTPASIPDDTLEKHTLRSESTYMLTVGDGSGILVFF 60
 QY 105 AVRPDGEIWMRHQTLPRD---NMDECTFEMMT-----NDVQPSFE 145
 DB 61 PGFP--GSVGAHYTLQSGSYQFDQ---MLTQNLPSVSYNCRVSRSLTVASSTL 113
 QY 146 MEGLOSINP--NAATGERRVLTGVPCKTGMQFDAETGEFLMARDTNYNMIESI---- 199
 DB 114 PGVYALNGTINAVT-----FGSLSEL---TDYSNGLSKATANIN 152
 QY 200 DENGIVYVNEAIIKEIDVEYDCPTFLGGRDWPASALNDPSGIYFIPLNNVC----- 252
 DB 153 DKIGNVIVGEGVTVLSPTSYDLSYRLGD-PIPAAGLDP-----KLMTCDSSDRPR 204
 QY 253 -YDMAAVDQ-EF-----TSMVYNTSNVTKLPPGKDM-GRIDAIDISTGRTLM- 298
 DB 205 VYTVTADEYQSSQLIPSGVKTTLFTANIDALTSVGGELIFSGVTHSIEVDVITYF 264
 QY 299 -----SVERAANYSPVLSGGLVFN--GCTDRYFRALSOETGETLMOTRLATVAS 348
 DB 265 IGFDTGEVTVKAVATDFGLTGTNNLVLPNLGGPTSEITQPTSMKLEVVYTKRGGT--A 322
 QY 349 GOAISYEVDGMQYVAIAGG 367
 DB 323 GDPISWTVSGTLAVTIYVG 341

RESULT 7
 PCT-US93-10541-2
 ; Sequence 2, Application PC/TUS9310541
 ; GENERAL INFORMATION:
 ; APPLICANT:
 ; TITLE OF INVENTION: TRANSGLUTAMINASE GENE
 ; NUMBER OF SEQUENCES: 5
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: RICHARDS, MEDLOCK & ANDREWS
 ; STREET: 1201 Elm Street, Suite 4500
 ; CITY: Dallas
 ; STATE: TX
 ; COUNTRY: USA
 ; ZIP: 75270-2197
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: PCT/US93/10541
 ; FILING DATE:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/126119
 ; FILING DATE: 23-SEP-1993
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/969702
 ; FILING DATE: 03-NOV-1992

ATTORNEY/AGENT INFORMATION:

```

NAME: Eugenia S. Hansen
REGISTRATION NUMBER: 31966
REFERENCE/DOCKET NUMBER: 33590C1PDPCT, OMRF B
TELECOMMUNICATION INFORMATION:
TELEPHONE: 214-939-4500
TELEFAX: 214-939-4600
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 548 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US93-10541-2

```

Query Match 4.7%; Score 95.5; DB 5; Length 548;
Best Local Similarity 21.4%; Pred No. 0.41;
Matches 77; Conservative 48; Mismatches 148; Indels 87; Gaps 19

```

QY      23  SGHSHAGGEELIMRNYLTIPRAGEDEDEFW-----GNDYERMMWIGAGOUITYIPVNLVH  16
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      17  NGRDHHTA-DLCREKLYVRRCO---PMLULTHEGRNYEA-----SVDSLFSFVVT-----  63
QY      77  YGSTRVAGASTOGCTGCGTGLYGNTRFEAVRPDGLIYWHQTLPRHNMDOECTFEEMVT  136
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      64  -----GAPSQA-----GTRAKPLDAVEEDMTATVYDO-----ODTLISQLT  105
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      137  NYDVOPSTREMGLOSINPNATGERR-----VLTGVPCKTIGTW-OPDAETGEF-L  185
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      106  T-----PANAPIGLYRLSLDASTGYOGSSEVLGHPILTFENAMPCPADAVYLDSEERREHYVL  161
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      186  WARDTNYO--NMIESIDENIYVNEBAILKELDVEYDCPEFL--GGRDPSAA-----  236
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      162  TQOGFTYOGSAKFKNIPWN--FCQFEGILDICLILIDNPKFLKNAGRDCSRSSPYV  219
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      237  -----INPDSGIFIPILNNVCYDMAAVDOFTSMDY--YNTSNYTKLEPGKDM  283
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      220  YGRVYSGMVNCONDQGVLLGRMDNNYGDGVSFMSIMASVDILIRMKNHGCRQVRYGCGCW  279
QY      284  GRIDLIDISTGRITLMSVRAAANYSPLYLSTGGVLFNGDIDRYR-----ALSGTGTETM  339
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      280  --FAAVACTYLCGCIPTVRVYTNNSAHONSNLLI-----EYFRNEBEGLOGKSEMIW  333

```

RESULT 8

US-08-216-276A-19
; Sequence 19, Application US/08216276A

; GENERAL INFORMATION:

1 APPLICANT: VAKHARIA, VIKRAM
2 APPLICANT: SNYDER, DAVID
3 TITLE OF INVENTION: SPECIFIC DNA AND RNA SEQUENCES
4 TITLE OF INVENTION: ASSOCIATED WITH US IDBV VARIANTS, VECTOR CARRYING DNA
5 TITLE OF INVENTION: SEQUENCES, HOST CARRYING CLONED VECTOR, DEDICED AMINO
6 TITLE OF INVENTION: SEQUENCES, VACCINE AND METHOD OF VACCINATION
7 NUMBER OF SEQUENCES: 34
8 CORRESPONDENCE ADDRESSES:

; COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.2
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/216,270A
 FILING DATE: 23-MAR-1994

CLASSIFICATION: 435
PRIOR APPLICATION DATA

```

1      PRIOR APPLICATION DATA:
2      APPLICATION NUMBER:  US 08/083,784
3      FILING DATE:  28-JUN-1993
4      PRIOR APPLICATION DATA:
5      APPLICATION NUMBER:  US 07/519,202
6      FILING DATE:  04-MAY-1990
7      PRIOR APPLICATION DATA:
8      APPLICATION NUMBER:  US 07/227,311
9      FILING DATE:  02-AUG-1988
10     ATTORNEY/AGENT INFORMATION:
11     NAME:  Kelber, Steven B.
12     REGISTRATION NUMBER:  30,073
13     REFERENCE/DOCKET NUMBER:  2747-054-27 CITE
14     TELECOMMUNICATION INFORMATION:
15     TELEPHONE:  (703) 413-3000
16     TELEFAX:  (703) 413-2220
17     TELEX:  248855 OPMT UR
18     INFORMATION FOR SEQ ID NO:  19:
19     SEQUENCE CHARACTERISTICS:
20     LENGTH:  1012 amino acids
21     TYPE:  amino acid
22     TOPOLOGY:  linear
23     MOLECULE TYPE:  protein
24     US-08-216-276A-19

```

Query Match	4.78;	Score 95;	DB 1;	Length 1012;
Best Local Similarity	23.88;	Pred. No. 1.2;		
Matches	82;	Conservative	43;	Mismatches 132; Indels 88; Gaps 20

```

OY      83 GPAS-----EORGRPGTIGTIRAVRBDGEIWMKQTRKD-----N 124
Db      24 GPASIPDDTLEKHTLRSESTSYNLTVCGTSGGLVIFEPGFP--GSIAGAHYTLLOSNNYK 81
OY      125 WDOECTFEMAYTNVDQVOSTEMEG--OSINPMATGERRLTGVPEKCTGMMQFDAET-- 181
Db      82 FQO-----MLTAQQLPSPYINCRVLSNLSYVRST---LPGSVYALNGII---NAVTF 129
OY      182 -GEFLMARDYNONNIESI---DENGIVYNEDAILKELDEYDVCPTEFLGDMPSAA 236
Db      130 QOSLSELTENVSYNGLSMATANMNDKICGNVLVGEQVYSLTPTSYDLGIVRBLGD--PIPAIG 188
OY      237 LNP-----DSG---IYEIFLANNVCYDMAVNDQFTSMDEVYNTSNV-----T 274
Db      189 LDPKVAATCDSSDRPRVYTI-----TAADDVQFSQ--YQYGGVYITLFSANIDAIT 238
OY      275 KLPCKDMIGRIDAIDISTGRTLSV--ERAAANYSPVLSGTGGVLENGGTDRY--FRAL 333
Db      239 SLVSQGEIVFKTSVSHLYGATIYLGPDGSAVITRAVAANG---LTTGTDNIMPENLV 295
OY      331 --SOEGETIMQTRLATVAS-----GGAISEYVQMGVYAIAG 367
Db      296 IPIPNELTQPTITSKLEIYTSKSGGEGEOMSMASAGSLAVYIHGG 340

```

RESULT 9

US-09-356-818A-2
; Sequence 2, Application US/09356818A

GENERAL INFORMATION:

? APPLICANT: Fraïf, Bassam M.
 ? TITLE OF INVENTION: Active Human Cellular Transglutaminase
 ? NUMBER OF SEQUENCES: 10
 ? CORRESPONDENCE ADDRESS:
 ? ADDRESSEE: Bassam M. Fraïf

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5, 1.44 Mb High Density Diskette

COMPUTER: IBM AT Compatible
OPERATING SYSTEM: Windows NT
SOFTWARE: WordPerfect 8.0 (saved in ASCII format)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/356,818A
FILING DATE: July 19, 1999
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME:
REGISTRATION NUMBER:
REFERENCE/DOCKET NUMBER: 2451.002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (405)652-7252
TELEFAX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 512
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-356-818A-2

Query Match 4.7%; Score 94.5; DB 4; Length 512;
Best Local Similarity 21.4%; Pred. No. 0.46;
Matches 77; Conservative 48; Mismatches 148; Indels 87; Gaps 19;

QY 23 SGHDSATGELMKNYFIIPRAGEGDETW-----GNDYFARMWTGAWGQITYPVTLVH 76
DB 17 NGRDHHTA-DLCREKLVRRGQ---PFWLTLHFGGRNYEA-----SVDSLTFSVYT----- 63
QY 77 YGTAIVGPASTGRTGGTLYGNTFPAVRPDTGEIVRHQTLPRDNDQECTFEEMWT 136
DB 64 -----GPAISOEA-----GTRARFPLRAVEEDMTATVVDQ-----QDCTLSLOLT 105
QY 137 NVDVQSTMEGLQSIINPAATGERR-----VLTGVPCKTGTMW-QFDAETGEF-L 185
DB 106 T-----PANAPIGLYRLLEASTGYOGSSFVLGHFILLFNAMCPADAVYLDSEERQEVYL 161
QY 186 WARDTNYQ---NMIESIDENGIVTNEAAILKELDEYDYVCTFL--GGRMPSAA----- 236
DB 162 TQGGFIYQSAKFINKIPWN--FGQFEDGILDICLLLDVNPFLKNAGHDCSRSSPY 219
QY 237 -----LNPDSGIYFIPLNNVCYDMAVADQEFSTMDV---YNTSNVTKLPPGKMI 283
DB 220 VGRVYSGMVCNDDQGVLLGRMDNNYGDGVSPMSWIGSVDIIRKKNHGCQRVKYGQCV 279
QY 284 GRIDAIDISTGRTLMSVERAANYSPLYSTGGVLFNGGTDYFR---ALSQETGETLM 339
DB 280 --FAAVACTVLRCLGIPTRVVTNYSADHONSLLI-----EYFNERFEGIEGDKSEMIW 332

RESULT 10
US-08-247-902A-2

; Sequence 2, Application US/08247902A
; Patent No. 5726051
; GENERAL INFORMATION:
; APPLICANT: Fraai, Bassam M
; APPLICANT: Birchbichler, Paul J
; APPLICANT: Patterson Jr., Manfred K
; APPLICANT: Gonzales, Robert A
; TITLE OF INVENTION: TRANSGLUTAMINASE GENE
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STIDLEY & AUSTIN
; STREET: 1201 Elm Street, Suite 4500
; CITY: Dallas
; STATE: TX
; COUNTRY: USA
; ZIP: 75270-2197
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/247,902A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 969702
FILING DATE: 03-NOV-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 126119
FILING DATE: 23-SEP-1993
ATTORNEY/AGENT INFORMATION:
NAME: Haite, John A.
REGISTRATION NUMBER: 37345
REFERENCE/DOCKET NUMBER: OMRF B-33590CIPCCIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: 214-981-3300
TELEFAX: 214-981-3400
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 548
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-247-902A-2

Query Match 4.6%; Score 93.5; DB 1; Length 548;
Best Local Similarity 21.4%; Pred. No. 0.65;
Matches 77; Conservative 48; Mismatches 148; Indels 87; Gaps 19;

QY 23 SGHDSATGELMKNYFIIPRAGEGDETW-----GNDYFARMWTGAWGQITYPVTLVH 76
DB 17 NGRDHHTA-DLCREKLVRRGQ---PFWLTLHFGGRNYEA-----SVDSLTFSVYT----- 63
QY 77 YGTAIVGPASTGRTGGTLYGNTFPAVRPDTGEIVRHQTLPRDNDQECTFEEMWT 136
DB 64 -----GPAISOEA-----GTRARFPLRAVEEDMTATVVDQ-----QDCTLSLOLT 105
QY 137 NVDVQSTMEGLQSIINPAATGERR-----VLTGVPCKTGTMW-QFDAETGEF-L 185
DB 106 T-----PANAPIGLYRLLEASTGYOGSSFVLGHFILLFNAMCPADAVYLDSEERQEVYL 161
QY 186 WARDTNYQ---NMIESIDENGIVTNEAAILKELDEYDYVCTFL--GGRMPSAA----- 236
DB 162 TQGGFIYQSAKFINKIPWN--FGQFEDGILDICLLLDVNPFLKNAGHDCSRSSPY 219
QY 237 -----LNPDSGIYFIPLNNVCYDMAVADQEFSTMDV---YNTSNVTKLPPGKMI 283
DB 220 VGRVYSGMVCNDDQGVLLGRMDNNYGDGVSPMSWIGSVDIIRKKNHGCQRVKYGQCV 279
QY 284 GRIDAIDISTGRTLMSVERAANYSPLYSTGGVLFNGGTDYFR---ALSQETGETLM 339
DB 280 --FAAVACTVLRCLGIPTRVVTNYSADHONSLLI-----EYFNERFEGIEGDKSEMIW 332

RESULT 11
US-08-971-782-4

; Sequence 4, Application US/08971782
; Patent No. 6043071
; GENERAL INFORMATION:
; APPLICANT: Wallis, Nicola G.
; APPLICANT: Shilling, Lisa K.
; APPLICANT: Jaworski, Deborah D.
; APPLICANT: Wang, Min
; APPLICANT: Mooney, Jeffrey L.
; APPLICANT: DeBouck, Christine M.
; APPLICANT: Zhong, Yi Yi
; TITLE OF INVENTION: No. 6043071el GIMU
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dechert Price & Rhoads

;; FILING DATE: 435
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 60/050,996
;; FILING DATE: 26-JUN-1997
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Dickinson, Todd Q
;; REGISTRATION NUMBER: 28,354
;; REFERENCE/DOCKET NUMBER: GM10024
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 215-994-2252
;; TELEFAX: 215-994-2222
;; TELEX:
;; INFORMATION FOR SEQ ID NO: 2:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 459 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; US-08-971-782-2

Query Match 4.6%; Score 93; DB 3; Length 459;
Best Local Similarity 26.3%; Pred. No. 0.55;
Matches 46; Conservative 26; Mismatches 69; Indels 34; Gaps 9;

QY 66 ITYPVTNLV--HYGSTAVGPASETORGTPGGLTYGNTREAVRPDGTGEIWMRHQTLPRD 123
DB 106 ITGESLKNLIDFHNHNKNAATILTAETDNFPG--YGR---IVRNDNAEVL--RMYEQKD 157
QY 124 NMDQCTFEEMVNTVDVQSTEM--EGLQSIINPNAATGERRV--LTGVPCKTG-----TM 174
DB 158 ATDFEKOIKEINTGYVDNERLFEALKNTNNAGGEYIITDVIGIFRETGEKVGAYTL 217
QY 175 WQFPAE-----TGEFLMARDNTYQMMISIDENGIVYVNDAILKEIDVE 219
DB 218 KDFESLGVNDVALATAESVMRRIRNHKHMV-----NGVSFVNPKATYIDIDVE 267

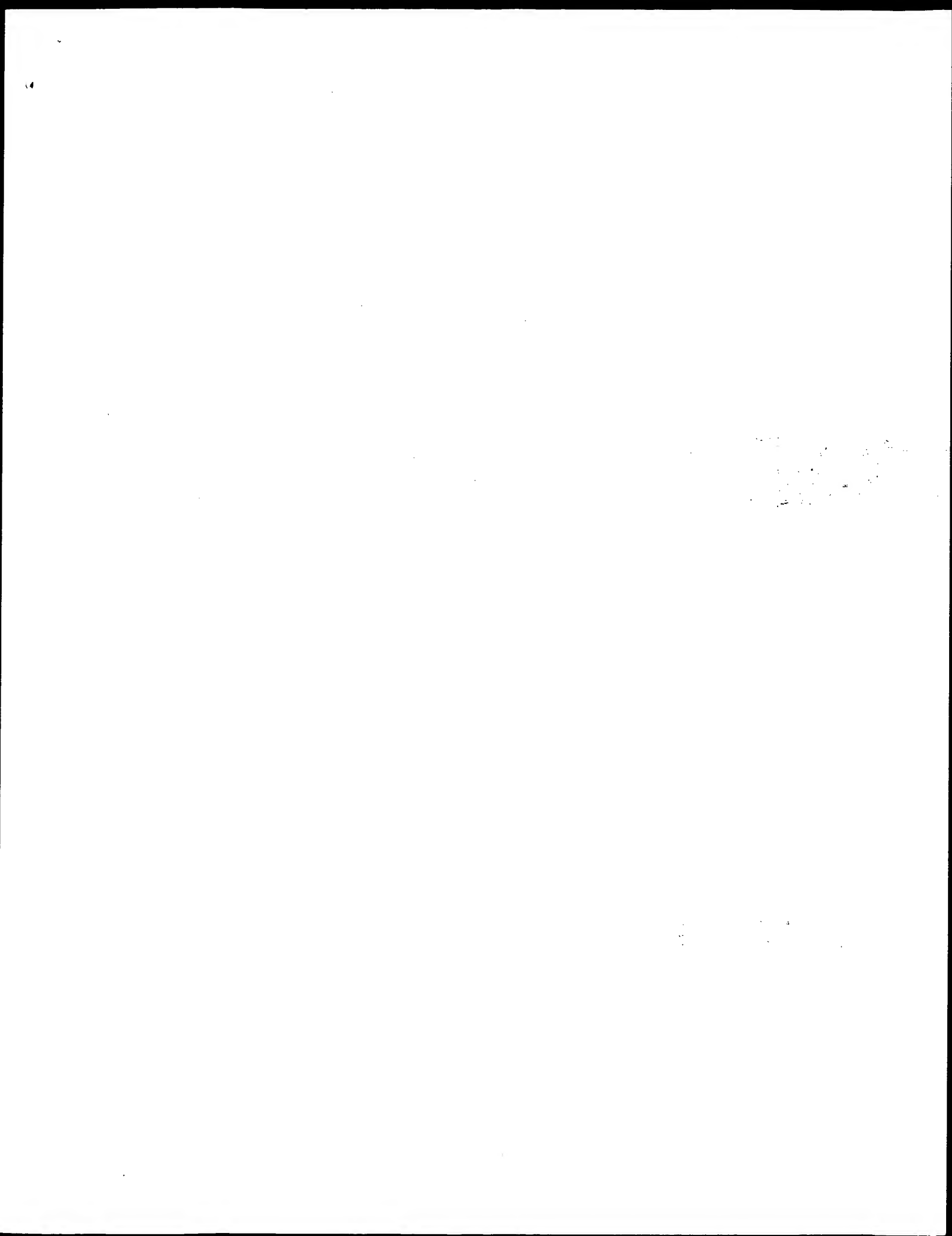
RESULT 14
US-09-309-026-2
Sequence 2, Application US/09309026
Patent No. 6204042
GENERAL INFORMATION:
APPLICANT: Wallis, Nicola G.
APPLICANT: Shilling, Lisa K.
APPLICANT: Jaworski, Deborah D.
APPLICANT: Wang, Min
APPLICANT: Mooney, Jeffrey L.
APPLICANT: Debouck, Christine M.
APPLICANT: Zhong, Yi Yi
TITLE OF INVENTION: No. 6204042el GIMU
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dechert Price & Rhoads
STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
CITY: Philadelphia
STATE: PA
COUNTRY: US
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/309,026
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/971,782
FILING DATE:
ATTORNEY/AGENT INFORMATION:

;; NAME: Dickinson, Todd Q
;; REGISTRATION NUMBER: 28,354
;; REFERENCE/DOCKET NUMBER: GM10024
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 215-994-2252
;; TELEFAX: 215-994-2222
;; TELEX:
;; INFORMATION FOR SEQ ID NO: 2:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 459 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; US-09-309-026-2

Query Match 4.6%; Score 93; DB 4; Length 459;
Best Local Similarity 26.3%; Pred. No. 0.55;
Matches 46; Conservative 26; Mismatches 69; Indels 34; Gaps 9;

QY 66 ITYPVTNLV--HYGSTAVGPASETORGTPGGLTYGNTREAVRPDGTGEIWMRHQTLPRD 123
DB 106 ITGESLKNLIDFHNHNKNAATILTAETDNFPG--YGR---IVRNDNAEVL--RMYEQKD 157
QY 124 NMDQCTFEEMVNTVDVQSTEM--EGLQSIINPNAATGERRV--LTGVPCKTG-----TM 174
DB 158 ATDFEKOIKEINTGYVDNERLFEALKNTNNAGGEYIITDVIGIFRETGEKVGAYTL 217
QY 175 WQFPAE-----TGEFLMARDNTYQMMISIDENGIVYVNDAILKEIDVE 219
DB 218 KDFESLGVNDVALATAESVMRRIRNHKHMV-----NGVSFVNPKATYIDIDVE 267

RESULT 15
US-08-883-534-3
Sequence 3, Application US/08883534
Patent No. 5846777
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Lal, Preeti
APPLICANT: Corley, Neil C.
TITLE OF INVENTION: TWO NEW WD-40 PROTEINS
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/883,534
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0332 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 606 amino acids



GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd

OM protein - protein search, using sw model

```
Run on:      May 24, 2002, 10:06:05 ; Search time 56.04 Seconds
              (without alignments)
              646.425 Million cell updates/sec
```

Title:	US-08-934-506A-5_COPY_180_556
Perfect score:	2032
Sequence:	1 IVANGVIAGSTQYSPFGC.....GMOYVAIAGGVSYSGLNS 377

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

```
Searched:      283138 seqs, 96089334 residues
Total number of hits satisfying chosen parameters: 283138
```

```

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%

```

```
Database : PIR_71:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*
```

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	374	18.4	601	2	E95863	alcohol dehydrogen
2	355.5	17.5	708	2	S52317	quinohemoprotein e
3	351.5	17.3	626	2	J00706	alcohol dehydrogen
4	338	16.6	742	2	J50336	alcohol dehydrogen
5	338	16.6	742	2	A49340	alcohol dehydrogen
6	335.5	16.5	573	2	S68591	methanol dehydroge
7	332.5	16.4	738	2	S14270	alcohol dehydrogen
8	308.5	15.2	623	2	B83390	alcohol dehydrogen
9	253	12.5	801	1	S00923	glucose dehydrogen
10	231.5	11.4	796	1	JV0107	glucose dehydrogen
11	231.5	11.4	796	2	H90644	glucose dehydrogen
12	231.5	11.4	796	2	B85455	glucose dehydrogen
13	222.5	10.9	796	2	AG0523	glucose dehydrogen
14	209.5	10.3	221	2	A41378	hypothetical protei
15	207	10.2	809	2	B98314	probable quinate d
16	207	10.2	809	2	A12968	hypothetical protei
17	204.5	10.1	809	2	A55547	guinate-shikimate
18	200.5	9.9	808	1	OPKEX	glucose dehydrogen
19	174	8.6	803	2	F83560	glucose dehydrogen
20	161.5	7.9	639	2	JC4861	polyvinyl alcohol
21	148.5	7.3	778	2	G98221	glucose dehydrogen
22	148.5	7.3	778	2	A13064	glucose dehydrogen
23	130.5	6.4	524	2	A82560	polyvinylalcohol d
24	123.5	6.0	668	2	C75264	probable serine/thr
25	119	5.9	407	2	H69064	serine/threonine p
26	118.5	5.8	385	2	A82284	conserved hypothet
27	118.5	5.8	799	2	T48869	serine/threonine P
28	118	5.8	827	2	F64512	hypothetical protei
29	117.5	5.8	253	2	AC0304	probable hemolysin

30	115	5.7	44.3	2	S23771	outer membrane por
31	114.5	5.6	1588	2	A86036	probable adhesin Z
32	114.5	5.6	1588	2	H91188	probable adhesin E
33	114	5.6	766	2	T33962	coatomer complex b
34	112	5.5	380	2	C83171	conserved hypoteth
35	112	5.5	839	2	D97013	probably secreted
36	110	5.4	352	2	AG0821	probable lipoprote
37	109.5	5.4	350	2	C70072	serine/threonine P
38	109.5	5.4	1197	2	D82696	hypothetical protel
39	108.5	5.3	583	2	SC36781	esterase D - Pseu
40	108	5.3	1526	2	AC2239	WP-40 repeat prote
41	106	5.2	776	2	SE5790	hypothetical prot
42	105.5	5.2	538	2	AF2099	hypothetical prot
43	105.5	5.2	593	2	S00128	protein G precurs
44	105	5.2	393	2	AD0350	probable lipoprote
45	103.5	5.1	473	2	D81322	outer membrane pro

ALIGNMENT:

```

RESULT      1
E95863      alcohol dehydrogenase (acceptor) (EC 1.1.99.8) large subunit [imported] - Sinorhizobi
C:Species:   Sinorhizobium meliloti
C:date:      24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 14-Sep-2001
C:Accession: E95863
R:Flinan,T.M.; Weldner,S.; Wong,K.; Buhrmester,J.; Chain,P.; Vorholter,F.J.; Her
Proc.Natl.Acad.Sci.U.S.A. 98, 9889-9894, 2001
A:title:     The complete sequence of the 1,683-kb pSymB megaplasmid from the N2-fixing e
A:Reference number: A95842; MUID:21396508; PMID:1148131
A:Accession: E95863
A:status:    preliminary
A:molecule type: DNA
A:Residues: 1-601 <KB>
A:Cross-references: GB:AL591985; PIDN:CAC48573.1; PID:g15140045; GSPDB:GN00167
A:Experimental source: strain 1021, megaplasmid pSYM B
R:Galibbert,F.; Flinan,T.M.; Long,S.R.; Punler,A.; Abola,P.; Ampe,F.; Barloy-Hubl
pela,D.; Chain,P.; Cowie,A.; Davis,R.W.; Dreano,S.; Federspiel,N.A.; Fisher,R.
L.; Hyman,R.W.; Jones,T.
Science 293, 668-672, 2001
A:Authors: Kahn,D.; Kahn,M.L.; Kalman,S.; Keating,D.H.; Kiss,E.; Komp,C.; Lelau
habault,P.; Vandenbol,M.; Vorholter,F.J.; Weldner,S.; Wells,D.H.; Wong,K.; Yeh,
A:title:     The composite genome of the legume symbiont Sinorhizobium meliloti.
A:Reference number: A96039; MUID:21368234; PMID:11474104
A:Contents:  annotation
C:Genetics:
A:Gene:      SMD20173
A:genome:    plasmid
C:Keywords:  oxidoreductase

Query Match          18.4%; Score 374; DB 2; Length 601;
Best Local Similarity 26.5%; Pred.No. 6,2e-21;
Matches 113; Conservative 66; Mismatches 161; Indels 86; Gaps 15;

QY      2 VANGYIVAGSTCGTQYSPFGCTFVSGHDSATGELMRNYFT-----PRAG 43
       | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
DB      181 VKDKILVIGISGEEFCVGRG-HVTAYSMADGKYLWRGYSMGPPSDPLIDPEKTHLGKPYGK 239
       ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      44 EHGDETFWGDYAEARMNTG---AMGOITYDPVTNLVHSYAVGPASFEORTPGGTLYGT 100
       || | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      240 DSGLTWTBED--ÖWKISGGTTWGMYSTDPEBNLVITGTGMPSTWNPTQR--PGDNRR-S 293
       || | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      101 MTRFAVRPDPTGEIWRHOQLTPRDNMDOECTFEEMVTVNDVPSTEMEGIASINPNAATGE 160
       || | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      294 MTFFARVDYTGAKAKLLYQMTPHEMDYDGVNEMILTEQQIDGK-----D 337
       || | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      161 RRVLTGVECKTGTMOQDAETGEEL-----WARDTYQNKKIE-----SI 199
       || | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      338 RLKLTHFD-RNGFYGTMDRTVGELLVAEKYDPTVMATEAVMDEPKSDYGRPOVVAOYST 396
       || | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      200 DENGLVYNEDAILKELDYEVDVCYTFILGGRDWPSAALNPDSGIFFIPLNNVCYDDMAVD 259

```

```

RESULT      2
552317
quinohemoprotein ethanol dehydrogenase (EC 1.1.99.-) type 1 precursor - Comamonas testos
C:Species: Comamonas testosteroni
C:Date: 08-May-1995 #sequenceRevision 21-Jul-1995 #text_change 02-Jun-2000
C:Accession: S62366; S62373; S65908; S52317
R:Stoorvogel, J.; Kiraayeld, D.E.; Van Suijs, C.A.; Jongejan, J.A.; de Vries, S.; Duine,
Eur. J. Biochem. 235, 690-698, 1996
A:Title: Characterization of the gene encoding quinohemoprotein ethanol dehydrogenase c
A:Reference number: S62366; MUID:96184549
A:Accession: S62366
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-708 <STO1>
A:Cross-references: EMBL:X81880; NID:g663195; PIDN:CA57464.1; PID:g663196
A:Accession: S62373
A:Status: Preliminary
A:Molecule type: protein
A:Residues: 32-54 <STO2>
R:de Jong, G.A.H.; Geerlof, A.; Stoervogel, J.; Jongejan, J.A.; de Vries, S.; Duine, J.A.
Eur. J. Biochem. 230, 899-905, 1995
A:Title: Quinohemoprotein ethanol dehydrogenase from Comamonas testosteroni. Purification
A:Reference number: S65908; MUID:95324580
A:Accession: S65908
A:Molecule type: protein
A:Residues: 32-50, 'X', 52-54, 477-483, 'X', 485-490 <DEJ>
A:Experimental source: ATCC 15667
C:Genetics:
C:Gene: qhcdh
C:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)
C:Keywords: oxidoreductase; quinoprotein
F:1-31/Domain: signal sequence #status predicted <SIG>
F:32-708/Product: quinohemoprotein ethanol dehydrogenase type 1 #status experimental <M
Query Match      17.5%; Score 355.5; DB 2; Length 708;
Best Local Similarity 26.7%; Pred. No.2,1e-19;
Matches 105; Conservative 55; Mismatches 180; Indels 49; Gaps 12;
QY      6 VVAGSTCYQSPGCFVSGHDSATGEELMRNYFP-----RAGEGDETWGNDYEA 56
      11: : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      208 VILKRGAEVGVRG-YITVADAEFGERKRMFVSVPDGPSPKFEDESKKRAART--DPSG 264
      11: : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      57 RMMTGA-----WQGITDPTNLVHGSTAVGPASEFGRTPGSTLIGTITRAVRPDTG 111
      11: : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      265 KMWEDAGGGGTMMWSMTFFDAELNTMYGTNGSPWSHRKRPKGGDNLIALSVALDPDTG 324
      11: : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      112 EIVRHOTTELRMMWDOCEFTFEMMYTNDVDPSTFEMEGLQSLNPAATGERVLTGVCPT 171
      11: : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      325 KYRMHNGEFGDMWDYISTQPMILADIKI-----AGPKKVLIIHAP-KN 367
      11: : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      172 GYMWQDAETGEFLMARDIYQNNIISIDENG-IVTVNEDALIKELDVEYDVCPTFLGGR 230
      11: : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      368 GFFPFLDRTNGKFIKSAKNFVPMWASGYDKHGKPIGI---AAARDGSKPDQAVPGPYGAH 424
      11: : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      231 DWFSALNPDSGIYIFPLNNVCYDMAAVDQ-EFT-----SMDYINISNVTKL-PPCK 280
      11: : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

RESULT      3
J00706      alcohol dehydrogenase (acceptor) (EC 1.1.99.8) alpha chain precursor - Methyllobacterium
N:Alternate names: methanol dehydrogenase 62K large chain
C:Species: Methyllobacterium extorquens
C>Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 08-Oct-1999
C:Accession: J00706; S07908
R:Anderson, D.J.; Morris, C.J.; Nunn, D.N.; Anthony, C.; Lidstrom, M.E.
Gene 90, 173-176, 1990
A>Title: Nucleotide sequence of the Methyllobacterium extorquens AM1 moxP and moxJ genes
A:Reference number: J00706; MUID:90337342
A:Accession: J00706
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-626 <AND>
A:Cross-references: GB:J31108; NID:g150017; PIDN:AAA25380.1; PID:g150018
A:Experimental source: strain AM1
R:Nunn, D.N.; Day, D.; Anthony, C.
Biochem. J. 260, 857-862, 1989
A>Title: The second subunit of methanol dehydrogenase of Methyllobacterium extorquens
A:Reference number: S04644; MUID:89350892
A:Accession: S07908
A:Molecule type: protein
A:Residues: 28-50,'XX','53 <NUN>
A>Note: The source is designated as Methyllobacterium extorquens AM1
C:Comment: This enzyme oxidizes methanol to formaldehyde.
C:Genetics:
A:Gene: moxP
C:Keywords: alcohol metabolism; oxidoreductase
F:1-27/Domain: signal sequence #status accepted <SIG>
F:28-626/Product: alcohol dehydrogenase (acceptor) #status experimental <MAT>

Query Match      17.3%; Score 351.5; DB 2; Length 626;
Best Local Similarity 26.2%; Pred. No. 3,6e-19;
Matches 106; Conservative 69; Mismatches 169; Indels 61; Gaps 14;

QY      1 IVANGVIVAGSTCOYSPFGCVSGHDSATGEELMRNYFI-----PRAGE 44
DB      191 VVKDKVILIGSSGAEIGVGRG-YLTAYDVKTGEQVWRAVATGPDKDLLASDFIKINPHYGQ 249
QY      45 E--GDEIWMGMDYIARWMTGA---WGOITPYVYNLVHYGSTAVGPASETORGTGGTLYG 99
DB      250 KGLGTGTWEED---AMKIGGGTMMGWYAYDPGNLILYFGTGNAPMNETMR--PGDNKW- 303
QY      100 TINTREAFARPTGIVWRHQTLPDNDDOECTFEMATYNDVVOESTENEGLOSINPAAATG 159
DB      304 TMTIFGRADPTGSAKRGYQKTPHDEMD-----YAGVAVMMLSECKD-----KDG 347
QY      160 ERRVLTGVPCKTGTMWOFDAETGEFLMAR---DTNYQNMIESIDENGIVTVNEDAILKET 216
DB      348 KARKLLTHPRNRGIVTYTLDRTDGALVSANKLIDDT--VNVFKSVDLTKGQVRNRPVETGRN 405
QY      217 D-VEIYVCPFLFGLGRWPSPSALNPPDSGIYFIPLANNCYDMMAAVDOEFTSMDYNNINSVTK 275
DB      406 DHIADKICPSAMYNHQGDHSYDPKRELFEMGINHICMWEDEPMLYRAGQEFVAGTLLNM 465
QY      276 LIPGK-----DMIGRIDLIDISTGRTLWSVEPRAAANYSPVLTGGGVLFNGTDTGYFRA 329
DB      466 YEPPKPDZRONYEGLGQIKIYNALITGDIKKWEKKEKRRFVWGGTMAATGADVLFYGTGLDYLKA 525
QY      330 LSGETGETLMQTRLATVASGAQISYVDQMOMVYVIAAGGVSVSGS 374

```



```

Db 125 ANGHLALDLAKTKINMEVEYCDPKVSTLTQAEPFAKDTVLKMGSCGALGVGAVNAPD 184
QY 27 SATGEELMR-----NYFIPRAGE--EGDETMGNDYEARMTGA---WGQ 65
Db 185 LKTELKMRFAATGSDSVRLAKDFNSANPHYCGFGLGTWEGD---AMKIGGTRMGW 241
QY 66 ITTDPVNLVHYGSTAVGPASETGRTGPG---TLVGTNTRFAVRPDTGELVWRHQL 120
Db 242 YADPKNTLFFYGGSGNPAPNMTETWR--PGDNKMTMTITWGRDL-----DTGMKKWGYOKT 293
QY 121 PRDMMDCGTEFEMVTVNDVQPSSTEMEGLOSINPNMATGRRVLTGVPCCTGTMNQPDAE 180
Db 294 PHDMWDAGVQWMLVD-----QPVN-----GAMTPDLSHIDNGILLYTLNRE 336
QY 181 TGEFLMARDTN-YONMIESIDENGIVYNEDAILKELDVE-YDVCPTFLGGRDMPAALN 238
Db 337 NGNLIVAEKVDPANVNRKDKLTGTPVRDPEFATRMDHGNICISAMFHNQVDSYD 396
QY 239 PDGSIYIPLNNVCYDMAVDQFTSMQVNTSNTKLP---PGKDMIGRIDALIDISTG 294
Db 397 PESHTILVAGNLHICMDPEPMLPYRAGQFVCGATLAMPGPNGPFRKEMQIRAFDLITG 456
QY 295 RTLMSVERAANYSPLYSTGGVILFNGCDYRFRALSOETGELTMQTRLATVAGQALSY 354
Db 457 KAKWTKKERRAANGCITLYTGGLVWATATLDGYLAKDNKDKELMFKMPSGGIGSPMTY 516
QY 355 EVDGMQVYALAGGVSYSG 374
Db 517 SFKGKQYI-----GSMWGVG 531

```

```

RESULT 7
S14270
alcohol dehydrogenase, membrane-bound (EC 1.1.-.-) 72k chain precursor - Acetobacter pol
C:Species: Acetobacter polyoxogenes
A:Variety: strain NB11028
C:Date: 21-Nov-1993 #sequence_revision 26-Jul-1996 #text_change 20-Jun-2000
C:Accession: S14270
R:Ramakli, T., Fukaya, M., Takemura, H., Tayama, K., Okumura, H., Kawamura, Y., Nishiyama
Biochim. Biophys. Acta 1088, 292-300, 1991
A:Title: Cloning and sequencing of the gene cluster encoding two subunits of membrane-bc
A:Reference number: S14270; MUID:91159482
A:Accession: S14270
A:Molecule type: DNA
A:Residues: 1-738 <TAM>
A:Cross-references: GB:D00635; NID:9216185; PIDN:BA00528.1; PID:9216186
A:Experimental source: strain NB11028
C:Complex: heterodimer of 72k and 44k chains
C:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)
C:Keywords: alcohol metabolism; membrane protein; oxidoreductase
F:1-33/Domain: signal sequence #status predicted <SIG>
F:36-738/Product: alcohol dehydrogenase 72k chain #status predicted <MAM>

```

```

Query Match 16.4%; Score 332.5; DB 2; Length 738;
Best Local Similarity 26.8%; Pred. No. 1.3e-17;
Matches 108; Conservative 53; Mismatches 179; Indels 63; Gaps 10;

QY 2 VANGVIVAGSTCOYSPFGC--FVSGHDSATGELMNRNYFIRAGEGD----- 47
Db 204 IAKGRVITII--GSEFEAGRGVSAFADATGKVDNRFFVTPPKNDPPDASVLMNKAY 261
QY 48 ETWGNDEARMTGA-----WGQITDPVNLVHYGSTAVGPASETGRTGPGTL 97
Db 262 QTWS-----PTGAWTRQGGGGTWDSTLYVDVADIVLYLGSGNSPMNKKYRSGKGN 314
QY 98 YGINTRFAYRPDTGELVWRHQLPRDMMDCGTEFEMVTVNDVQPSSTEMEGLOSINPAA 157
Db 315 LFLGSIYALKPTEGTYVWHFOETPMQDMFTSDQIMTLDLP----- 357
QY 158 TGERRVLTGVPCCTGTMNQPDAEETGELMARDTNQNMIESID-ENGIVYNEDAILKEL 216

```

```

Db 358 NGETRHVIVHARKNGFEYIIDAKTGEFISGKNVYVNMASGLDPKTRPIYNPDALYTLT 417
QY 217 DVEYDVCPTFLGGRDMPAALNPDGSIYIPLNNVCYDMAVDQFT-SMDVYNTS--N 272
Db 418 GKEWYGIPEDDLGNHFAAFESPCKTGLYIIPAOQVPLTYNOVGFTIHPHSMNIGLDMN 477
QY 273 VKRLPG-----KDMIGRIDALIDISTGRTLMSVERAANYSPLYSTGGVILFNGCTD 324
Db 478 KVGIDPSPEAKQAFVKDLKGLVAMDPOKQAEAMRVHDHKGPMWNGILATGDDLLFQGLAN 537
QY 325 KYFRALSOETGELTMQTRLATVAGQALSYEVDGMQVYALAGG 367
Db 538 GEFHAYDATNGSDLFHFAADSGIAPPVYTLANGKQYAVEVG 580

```

```

RESULT 8
B83399
quinoprotein alcohol dehydrogenase PA1982 [Imported] - Pseudomonas aeruginosa (strain
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: B83399
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warriner, P.; Hickey, M.J.;
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; L
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pa
A:Reference number: A82950; MUID:20437337
A:Accession: B83399
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-623 <STO>
A:Cross-references: GB:AE004624; GB:AE004091; NID:99447973; PIDN:AA005370.1; GSPDB:GN
A:Experimental source: strain PA01
C:Genetics:
A:Gene: exaB; PA1982

```

```

Query Match 15.2%; Score 308.5; DB 2; Length 623;
Best Local Similarity 22.9%; Pred. No. 7.3e-16;
Matches 88; Conservative 68; Mismatches 175; Indels 53; Gaps 10;

QY 26 DSATGEELMRNYFIR-----AGEGDETMGNDYEA-----RMATGA---WG 64
Db 224 DPTGGEELIMRPFVEGHRGLNGKDSITVGVKAPSWPDDNSPTGVESSHGGAPWQ 283
QY 65 QITDPVNLVHYGSTAVGPASETGRTGPGS-----TLVGTNTRFAVRPDTGELVWRHQ 118
Db 284 SASPEAETNTIIVGAGNPGPNWTAKAGGNPHDYSLY-TSGQGVDPSSGGEVWFYQ 342
QY 119 TLPRDMMDCGTEFEMVTVNDVQPSSTEMEGLOSINPN-----AATGERRVLTGVPCCTGT 173
Db 343 HTPNDAMDFSGNNELVLEFDYKAKGKIVKATAHADRNGFFVYVDRSGKIQNMFPEVDNI 402
QY 174 MW--QFAETGEFLMARDTNQNMIESIDENGIVTVNEDAILKELDVEYVCPFLAGRD 231
Db 403 TWASHIIDLKTG-----RVERESQRPPLPEPG-----QKHGRAVEYSPFLGKN 447
QY 232 WPSAALNPDGSIYIPLNNVCYDMAVDQFTSMQVNTSNTKLPKPGKMIGRIDALID 291
Db 448 WNPAAVYSDTDLFFVYPAHMKEDYVTEVSTYKSAVLGAGFRKRRKRYDDHVSLSRAMDP 507
QY 292 STGRTLMSVERAANYSPLYSTGGVILFNGCTDYRFRALSOETGELTMQTRLATVAGQA 351
Db 508 VSGKVVWEHKKHPLMLMGVILATAGNLVFTGCGDYFAFAFAKSGKELMKRTQTSGLVSP 567
QY 352 ISYEVDGMQVYALAGGVSYSGSL 375
Db 568 ITWBQDGEQYLGVY---VGYGAV 588

```

```

RESULT 9
S00943
glucose dehydrogenase (pyrroloquinoline-quinone) (EC 1.1.99.17) - Acinetobacter calco

```



```

Db      559  CRVMEHQMRKEGLETFPPSEQGLTVFPGNLGMFEWNGCISVDPEPREVALNPALMPLEVSRLI   618
               : : : : : | : : : : :
Oy      240  -----DSGI-----YFIPLNNCCYDMANADQETSNDVNTSNV   273
               : : : : : | : : : : :
Db      619  PRGPGNPMPOPKDANKGTCTGESGIQPOLGVVPYGVTLN-----PFLS-----P   659
               : : : : : | : : : : :
Oy      274  TLKPPGKMIGRIDAIDISTGRITLMSVERAAANY-----PV-----LSTG   314
               : : : : : | : : : : :
Db      660  FELPCQOPAMGYISALDKLTNEVWKRRIGTPQDSMPPEMPPVPVFNNGMPTLGSPISTA   719
               : : : : : | : : : : :
Oy      315  GGVLENCGI-DKYFAISOETGETLMOTRLATVASGA--ISEYENDQMAYAI-AGGSVS   370
               : : : : : | : : : : :
Db      720  GNVLFIATADNYLTRAYNMNSGEKLMOGRLP--AGQAATPMTYEVNGKOYVVISAGHGHS   777
               : : : : : | : : : : :
Oy      371  YGS   373
               : : : : :
Db      778  FGT   780

RESULT  12
H85495
glucose dehydrogenase [imported] - Escherichia coli (strain O157:H7, substrain EDL9333)
C:Species: Escherichia coli
C:Date: 16-Feb-2001 #sequence.revision 16-Feb-2001 #text.change 14-Sep-2001
C:Accession: H85495
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; May
Miller, L.; Grobbeck E.J.; Davis, N.W.; Lim, A.; DiMaio, P.A.; Potamousis, K.; Brodeur

```

A:Title: genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: H85495
A>Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-796 <STO>
A:Cross-references: GB:AE005174; NID:g12512839; PIDN:AAGS4428.1; GSPDB:GN00145; UWGP:
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: gcd
C:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)

Query Match 11.4%; Score 231.5; DB 2; Length 796;
Best Local Similarity 23.4%; Pred. No. 9e-10;
Matches 113; Conservative 53; Mismatches 154; Indels 163; Gaps 23;

OY 6 VIVAGS--TCQYS--PFCGFSGHDSATGCELMRYFLPRA-----GEGDETGWGDYE 55
:::||||| : | : | | | | | | | | | | | | | | | | | | | | | | | | |
Db 346 IVMAGSVTDNSTRETSVGINGEDVNTGCELIMA--FDPGAKDPAIMPISDEHTFFNS--- 400

```

Db 401 -----PNSWAPAAADAKIDLVLPMGVTPPTIWGNGNRPEQERAAST-ILALNATGGLAW 455
QY 116 RHQTLPRDNNDQECTEEMATVNDVQSTMEGLQSLNPAAATGERTVLGVGCKGTW 175
Db 456 SYCTVHHDLMDMD-----LPAQPT-----LADITVA---GQKPYIYAAKAGNIF 498
QY 176 QPAETGEEL-----WARDTNYQNNIESIDENGT-----VTVEDAI 212
Db 499 VLDNRNSELVPAPEKVPQGAAGDVTPTQPSSEISFRTKDLGADMGATMEDOLV 558
QY 213 LKEL--DVEYD-----VCPTFLGGRMPSAALNP----- 239
Db 559 CRVAFHQMRYEGLFTPPSEQGLVFPNGLGMFEMWGASVDNPKEVALANPMALPEVSKLL 618
QY 240 -----DSGI-----YFPLNANVCYDMAAVQDEFTSMDYVITSNV 273
Db 619 PRGFQGNMEQPKAKGTGSGIGIOPQYGVPGYTLN-----PPLS-----P 659
QY 274 TKLPQKMDIGRIDALIDISTGRTLSWVERAANYS-----PV-----LSTG 314
Db 660 FGLPQCPQPMAGVTSALDLTLTNEVYMKRKRIQTPQDSNPFPMYVPPNMGMPMLGIGISA 719

```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2002, 10:15:50 ; Search time 27.86 Seconds
(without alignments)
523.951 Million cell updates/sec

Title: US-08-934-506a-5_COPY_180_556

Perfect score: 2032
Sequence: 1 IVANGYIVAGSTGQYSPFGC.....GMOYVIAAGGVSYSGSLNS 377

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	361	17.8	600	1	XOFE_PARDE
2	351.5	17.3	626	1	DMH1_METEX
3	341.5	16.8	626	1	DMH1_METOR
4	340	16.7	631	1	DMH1_PAREB
5	338	16.6	739	1	DHET_ACCEU
6	338	16.6	742	1	DHET_ACCEC
7	332.5	16.4	571	1	DMH1_METME
8	332.5	16.4	738	1	DMH1_METME
9	308.5	15.2	623	1	EXAA_PSEAB
10	298.5	14.7	757	1	DHET_GLIOX
11	293	14.5	801	1	DHGA_ACICTA
12	231.5	11.4	796	1	DHG_ECOLI
13	204.5	10.1	809	1	OUIA_ACICA
14	200.5	9.9	808	1	DHG_GLIOX
15	180.5	8.9	790	1	OUIA_XANCI
16	118.5	5.8	799	1	ASCK_STROA
17	118	5.8	827	1	YK23_METTA
18	115.5	5.7	353	1	YKAL_BACSU
19	115.5	5.7	354	1	CADN_MOUSB
20	115	5.7	443	1	PORD_PSEAB
21	114	5.6	796	1	COPP_SCHPO
22	110.5	5.4	3317	1	CADN_RAT
23	108.5	5.3	3354	1	CADN_HUMAN
24	105.5	5.2	593	1	SPG2_STRSP
25	102	5.0	790	1	PLMN_PIG
26	102	5.0	1645	1	OMPB_RICVY
27	101	5.0	392	1	YFGL_ECOLI
28	101	5.0	1012	1	PORG_IBDVO
29	99.5	4.9	466	1	MMO8_RAT
30	99.5	4.9	807	1	AFSK_STRGR
31	98.5	4.8	1016	1	PMPH_CHLTR
32	96.5	4.7	465	1	MMO8_MOUSE
33	95.5	4.7	746	1	FEPA_ECOLI

34	95	4.7	570	1	SYG_PYRHO
35	94.5	4.7	746	1	FEPA_PSEAB
36	94	4.6	546	1	CH60_BURCE
37	94	4.6	546	1	CH60_BURVI
38	93.5	4.6	545	1	CH63_BRAVA
39	93.5	4.6	687	1	TGIC_HUMAN
40	93.5	4.6	2248	1	CVAL_DROME
41	93	4.6	537	1	ANPC_BOVIN
42	93	4.6	1365	1	GTF5_STROD
43	92.5	4.5	954	1	FLRY_CAUCR
44	91.5	4.5	760	1	MYTE_MYCLE
45	91	4.5	504	1	VL2_HPV63

ALIGNMENTS

RESULT	1	STANDARD	PRT	600 AA
XOFE_PARDE				
ID	XOFE_PARDE			
AC	P29958;			
DT	01-APR-1993 (Rel. 25, Created)			
DT	01-OCT-1996 (Rel. 34, Last sequence update)			
DT	15-DEC-1998 (Rel. 37, Last annotation update)			
DE	Putative dehydrogenase XOFE precursor (EC 1.1.99.-)			
GN	XOFE			
OS	Paracoccus denitrificans.			
OC	Bacteria; Proteobacteria; alpha subdivision; Rhodobacter group;			
OX	NCBI_TaxID=266;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Harms N.;			
RL	Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.			
RN	[2]			
RP	SEQUENCE OF 380-600 FROM N.A.			
RC	STRAIN=PD 1235;			
RC	MEDLINE=92041583; Pubmed=1657873;			
RA	Ras J., Reijnders W.N.M., van Spanning R.J.M., Harms N., Oltmann L.F.,			
RA	Scouthamer A.H.;			
RT	"Isolation, sequencing, and mutagenesis of the gene encoding			
RT	cytochrome c553 of Paracoccus denitrificans and characterization of			
RL	J. Bacteriol. 173:6971-6979(1991).			
CC	-1- COFACTOR: POO (BY SIMILARITY).			
CC	-1- SIMILARITY: BELONGS TO THE BACTERIAL POO DEHYDROGENASE FAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; U34346; AAC44555.1; -			
DR	EMBL; M75583; AAA25574.1; -			
DR	PIR; A41378; A41378.			
DR	HSSP; P38539; AAAH.			
DR	InterPro; IPR002372; Bac_POO_repeat.			
DR	Pfam; PF01011; Bacterial_POO; 7.			
KW	Oxidoreductase; POO; Signal.			
FT	SIGNAL	1	21	POTENTIAL.
FT	CHAIN	22	600	PUTATIVE DEHYDROGENASE XOFE.
FT	ACT_SITE	318	318	BASE (POTENTIAL).
SO	SEQUENCE	600 AA;	65159 MW;	DCA996PFBCCCA3CE CRC64;
QY	2	VANGYIVAGSTGQYSPFGFVSGHDSA-----TGEELMWNRYI-----		39
		Query Match	17.8%; Score 361; DB 1; Length 600;	
		Best Local Similarity	27.2%; Pred. No. 1.5e-20;	
		Matches	115; Conservative 61; Mismatches 167; Indels 80; Gaps 16;	

```

Db 180 VKDVLVIGSGEGG-----VRGMMTALNTLDSSEAKWASTGPDEELVDPETTHLCK 234
QY 40 PRAGEEDDETNGNDYERMMTGA---WGQITTPVNTLVHYSTANGPASETORPGT 96
Db 235 PIGADSSLSNMBD---QMOIGGTTWGFSTYDPLNLVYGGNPNSTWMPQR--PGDN 289
QY 97 LYGNTFFAVRPDTGELVWRHQTLPKDMNDDECTFEKMTN--VDYOPSTMEGLQSI 154
Db 290 KM-SMTIARDADTGMAKMFYQMTPHDEMDYDGVNMIILNQYDQ----- 335
QY 155 NAATGERVLTGVPCKTGTMQDAETGEFLVARDTNYOMIE-----SIDENGIYVNE 209
Db 336 -----EKKLTTHFD-RNGLAITYDRETBELVAE--KYDPVWMTTGVMDPNS-ETYGR 386
QY 210 DAILEKEDYED-----VCPTEFGGRDWPSSALNPDGSIYFIPLNWCVDMADVD 261
Db 387 PAVVAETSTQNGEDENTGVCAALGTDOOPAFSPKTNLYFTVTHVCMDEPERVA 446
QY 262 FTSMDVY--NTSNVTKLPCKDMIGRIDAIDISTGRTLMSVERAANYSPVLSTGGVLF 319
Db 447 YTAGOPVYVGTATLSMYPAPNSHGMNFIAWHNTTGEIKMSVPEQSPVSGALATAGDYF 506
QY 320 NGSTIDYFRALSGTGETIMQTLAVASGQALSYVDQMUYALAG-----GVSYSGS 374
Db 507 YGTEGTLKPVDAQGTSEELKFKTPSGIIGNWYTERHGKQYVILSGVGWAGIGLAG 566
QY 375 LNS 377
Db 567 LTN 569

RESULT 2
DHML_METEX STANDARD: PRT: 626 AA.
ID DHML_METEX
AC P16027;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Methanol dehydrogenase subunit 1 precursor (EC 1.1.99.8) (MDH large
GN alpha subunit) (MDH).
OS MOXF.
OS Methylobacterium extorquens.
OC Bacteria: Proteobacteria; alpha subdivision: Rhizobiales group;
OC Methylobacterium group; Methylobacterium.
OX NCBI_taxid=408;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=AM1 / NCIB 9133;
RX MEDLINE=90337342; PubMed=2116368;
RA Anderson D.J., Morris C.J., Nunn D.N., Anthony C., Lidstrom M.E.;
RT "Nucleotide sequence of the Methylobacterium extorquens AM1 moxf and
RL moxf genes involved in methanol oxidation.";
RL Gene 90:173-176(1990).
RN [2]
RN SEQUENCE OF 28-53.
RC STRAIN=AM1 / NCIB 9133;
RX MEDLINE=89350892; PubMed=2504152;
RA Nunn D.N., Day D., Anthony C.;
RT "The second subunit of methanol dehydrogenase of Methylobacterium
RL extorquens AM1.";
RL Biochem. J. 260:857-862(1989).
RN [3]
RN DISULPHIDE BONDS.
RX MEDLINE=95384759; PubMed=7656012;
RA Blake C.C.F., Ghosh M., Harlos K., Avezoux A., Anthony C.;
RT "The active site of methanol dehydrogenase contains a disulphide
RL bridge between adjacent cysteine residues.";
RL Nat. Struct. Biol. 1:102-105(1994).
RN [4]
RN X-RAY CRYSTALLOGRAPHY (1.94 ANGSTROMS).
RX MEDLINE=95253818; PubMed=7735834;
RA Ghosh M., Anthony C., Harlos K., Goodwin M.G., Blake C.;

```

```

RT "The refined structure of the quinoprotein methanol dehydrogenase
RT from Methylobacterium extorquens at 1.94 A.";
RL Structure 3:177-187(1995).
CC -1- CATALYTIC ACTIVITY: A primary alcohol + acceptor - an aldehyde +
CC reduced acceptor.
CC -1- COFACTOR: POO.
CC -1- SUBUNIT: THE HOLOENZYME MDH HAS AN ALPHA-2/BETA-2 CONFIGURATION
CC OF A LARGE ALPHA AND A SMALL BETA SUBUNIT.
CC -1- SUBCELLULAR LOCATION: PERIPLASMIC, BUT ASSOCIATED WITH INNER
CC MEMBRANE.
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POO DEHYDROGENASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: M31108; AAA25380.1; -.
DR PIR: S07908; J07908.
DR PIR: JQ0706; JQ0706.
DR HSSP: P38539; 4AAB.
DR InterPro: IPR001479; Bac_POO.
DR InterPro: IPR002372; Bac_POO_repeat.
DR Pfam: PF01011; Bacterial_POO_7.
DR PROSITE: PS00363; BACTERIAL_POO_1; 1.
DR PROSITE: PS00364; BACTERIAL_POO_2; 1.
DR KW Oxidoreductase; POO; Signal; Methanol utilization; Periplasmic.
FT CHAIN 28 626 METHANOL DEHYDROGENASE SUBUNIT 1.
FT DISULFID 130 131
FT FT 413 442
FT DISULFID 413 442
FT ACT_SITE 330 330
FT FT 330 330
SQ SEQUENCE 626 AA; 68434 MW; 64988DDAFDZAD34C CRC64;

Query Match 17.3%; Score 351.5; DB 1; Length 626;
Best Local Similarity 26.2%; Pred. No. 8.8e-20;
Matches 106; Conservative 69; Mismatches 169; Indels 61; Gaps 14;

QY 1 IVANGIVASTCOYSPFGCFVSGHDSATGEELMRNYFI-----PAGE 44
Db 191 VKDKVILIGSSGAEALGYRG-YLTAYDVKTGEQVWRAVATGPDKDLLASDFNIKKNHYGQ 249
QY 45 E--GDETGWGNDYERMMTGA---WGQITTPVNTLVHYSTANGPASETORPGTLYG 99
Db 250 KGLSTGWEGD---AMKIGGGTNGWYAYADPGTNLYFTGNDPAPNNTMR--PGDNKN- 303
QY 100 TNRFAVRPDTGELVWRHQTLPKDMNDDECTFEKMTN--VDYOPSTMEGLQSIINPAATG 159
Db 304 TMTIFGADATGKAKTKYQKTPHDEMD-----YAGVAVMLLSOKD-----KKG 347
QY 160 ERRVLGVPCKTGTMQDAETGEFLMAR--DTNYOMIESIDENGIYVEDALIKEL 216
Db 348 KARKLTHPDRNGIYVTLRTDGAIVASANKLDDF--VNFKSVDLKTGGPVRDPERGTGA 405
QY 217 D-VEYDVCPEFLGGRDWPSSALNPDGSIYFIPLNWCVDMADVDQFTSMYDYNSTYK 275
Db 406 DHIKADICPSAMHYNHQGHDSYDPKRELFPMGNHNCIMWEPFNLRYRAGQEFFVATLWM 465
QY 276 LPPGK-----DMIGRIDAIDISTGRTLMSVERAANYSPVLSTGGVLFNGCTDRYRA 329
Db 466 YRPKDRQNYEBGLQIKAVNATGTYKKMKERRAVWNGTMAATGDLVFTYGLDGIYLA 525
QY 330 LSGTGETIMQTLAVASGQALSYVDQMUYALAGGVSYSGS 374
Db 526 RDSDTGDLMLKFKIPSGAIGPMTYTHKGTQYVAI-----YGVGV 565

RESULT 3
DHML_METOR

```

```

ID DHM1_METOR STANDARD: PRT; 626 AA.
AC P15279;
DE 01-APR-1990 (Rel. 14, Created)
DE 01-APR-1990 (Rel. 14, Last sequence update)
DE 01-MAR-2002 (Rel. 41, Last annotation update)
DE Methanol dehydrogenase subunit 1 precursor (EC 1.1.99.8) (MDH large
DE alpha subunit) (MEDH).
GN MOXF.
OS Methylobacterium organophilum XX.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Methylobacterium group; Methylobacterium.
OX NCBI_TaxID=410;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 29-43.
RC STRAIN-ATCC 27866 / DSM 760 / NCIB 11278;
RX MEDLINE=89008094; PubMed=2459109;
RA MacMillin S.M., Hanson R.S.;
RT "Nucleotide sequence and transcriptional start site of the
RT Methylobacterium organophilum XX methanol dehydrogenase structural
RT gene."
RL J. Bacteriol. 170:4739-4747(1988).
CC -I- CATALYTIC ACTIVITY: A primary alcohol + acceptor = an aldehyde +
CC reduced acceptor.
CC -I- COFACTOR: PQQ.
CC -I- SUBUNIT: THE HOLOENZYME MDH HAS AN ALPHA-2/BETA-2 CONFIGURATION
CC OF A LARGE ALPHA AND A SMALL BETA SUBUNIT.
CC -I- SUBCELLULAR LOCATION: PERIPLASMIC, BUT ASSOCIATED WITH INNER
CC MEMBRANE.
CC -I- SIMILARITY: BELONGS TO THE BACTERIAL PQQ DEHYDROGENASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: M22629; AAA50289.1; -.
CC DR HSSP; P38539; 4AAH.
CC DR InterPro: IPR001479; Bac_PQQ.
CC DR InterPro: IPR002372; Bac_PQQ_repeat.
CC DR Pfam: PF01011; Bacterial_PQQ_7.
CC DR PROSITE: PS00363; BACTERIAL_PQQ_1; 1.
CC DR PROSITE: PS00364; BACTERIAL_PQQ_2; 1.
CC KM Oxidoreductase; PQQ; Signal; Methanol utilization; Periplasmic.
CC FT SIGNAL 1 28
CC FT CHAIN 29 626 METHANOL DEHYDROGENASE SUBUNIT 1.
CC FT DISULFID 130 131 BY SIMILARITY.
CC FT DISULFID 413 442 BY SIMILARITY.
CC FT ACT_SITE 330 330 BASE (POTENTIAL).
CC SEQUENCE 626 AA; 68677 MW; 8768F6B8371E5DF CRC64;

```

Query Match 16.8%; Score 341.5; DB 1; Length 626;
 Best Local Similarity 26.2%; Pred. No. 5.2e-19;
 Matches 106; Conservative 67; Mismatches 171; Indels 61; Gaps 14;

```

QY 1 IVANGVIAGSTGQSPFGVSGHSATGEELRNFTI-----PAGE 44
DE 191 VVADKVIIGSSGALGVRG-YILAYDYKTGGQVRAVATGPDLLADFNKNAHYGQ 249
QY 45 E--GDETWGNDYBARMTGA--WGQITVDPVNLVHYGSTAVGPASETGRTGGTLYG 99
DB 250 KGLGTATWED--AWKIGGATNMGWAYADPGTNLIYFGGNPAWMEIWR--PGDKW- 303
QY 100 TINTREAVRPTGTEIWRHQLPRDNDQECTFEMNTNVOPSTMEGLOSINPNAATG 159
DB 304 TMTIFGDAOTGAKFGYQATPHDEWDYAGVNM-----PSEQKD-----KDG 347
QY 160 ERYVLTGVPCKTMMQFDATETGEFLMAR--DTNYONMTESIDENGIVVNEADAILKEL 216
DB 348 KTKRLILHPDRNGLIVYTLDTGDAIVASANKLDTI--VNVKTYDLKTKGQVVRDEYCTRM 405

```

```

QY 217 D-VEYDVCFTEIAGRDWPSAALNPDSCITFIPLANVCYDMAVAODEFTSMYNTNSVTK 275
DE 406 DILANDVCPASGANGYHNOGHDSYDKRELFPMGINHICMDEPFMLPYRAGQFVGATLNN 465
QY 276 LPPGR-----MIGRIDIDISTGRTLMSVERAANYSPLVSTGGGVFNNGTDRFYRA 329
DB 466 YGPPGGRONRYEGLGQIKYVNAITGSGYKWEKMERFVWGGTATAGDLVYFGILDTLKA 525
QY 330 LSGETGELMOTRLATVAGSGAISEVDGMQVYALAGGVSYGSG 374
DB 526 RSDITDILLMKRKIPSGALGYPMFTYHKQYVAL-----YGVG 565

```

Query Match 16.7%; Score 340; DB 1; Length 631;
 Best Local Similarity 25.0%; Pred. No. 6.9e-19;
 Matches 107; Conservative 65; Mismatches 166; Indels 90; Gaps 15;

```

QY 1 IVANGVIAGSTGQSPF-----GC-----FVSGHSATGEELRNFTI----- 39

```

```

Db 178 IMENSDIKVSTLTIAPIYKIDLVYSSGSAELGVRGVTAYDVKSGEMRRARATGPDE 237
OY 40 -----PRAGEE--GDETGNDEYARMTGA---WGQITYPVTNLVHYGSTAV 82
Db 238 ELLIAEDFNAPNPHYGKNIQLETWECD---AMKIGGTMNGWAVAYDEVDLFFYSSGNP 294
OY 83 GPASEPQRTGPGGLYLTNTREAVRPTGELVWHQTLPRDNDWQDETFEMATNVDP 142
Db 295 APNMEYTR--PGDNKW-TMAIWGREATTGEAKFYQKTPHDEM-----YAGVVM 343
OY 143 STEMEGQSINPNATGERVRLTGVPCKTGMQFQDAETGPEFLWARDNTYQNMIESIDE- 201
Db 344 LSEBEDQ-----GQMKRLTHPRNGIYVTLDRNTGDLISA-----DKMDDT 386
OY 202 -NGIVYNEDAILKELDEY-----DYCPTEFLGRDMPASALNDPSGIYPIPLNV 251
Db 387 VNVWKEVQDLTGLEVRDPERGTRMDHKARDICPSAMGYHNGHSDYPERKVFILGNIH 446
OY 252 CYDMAVADQFTSMQVNTSNVTKLPKDKM-----IGRIDAIDISTGRTLSYERAAN 306
Db 447 CMCWEPMLYRAGQFVGTALTMYPGPKATPATERAGQIKAYDAISGEMKMEKERSV 506
OY 307 YSPVLSTGGGVLTNGTDRFRLALSOETGETIMQTRLATYASGQALSYEVDGMQYVIA 366
Db 507 WGGTMTAGGLTYVILDGFIKARDSDPTGDLMKFLPSGVIGHPMTYKHGROYVAI-- 564
OY 367 GGVSYSG 374
Db 565 ---MYGVG 569

RESULT 5
DHET_ACEEU STANDARD; PRT; 739 AA.
AC 044002; 007952;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Alcohol dehydrogenase [acceptor] precursor (EC 1.1.99.8).
GN ADH.
OS Acetobacter europaeus.
OC Bacteria; Proteobacteria; alpha subdivision; Acetobacteraceae.
OX NCBI_TaxID=33995;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DES11 / DSM 6160;
RA Submitted (NOV-1996) to the EMBL/Genbank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: A primary alcohol + acceptor = an aldehyde +
CC reduced acceptor.
CC -1- COFACTOR: POQ AND HEME (BY SIMILARITY).
CC -1- SUBUNIT: HETEROTETRAMER (DEHYDROGENASE, CYTOCHROME, AND TWO
CC SMALLER UNKNOWN SUBUNITS) THAT FORMS THE ALCOHOL DEHYDROGENASE
CC COMPLEX (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: MEMBRANE-BOUND, FACING THE PERIPLASMIC
CC SPACE (POTENTIAL).
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POQ DEHYDROGENASE FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X82894; CAA58066.1; -
CC EMBL; Y09480; CAA70688.1; -
CC HSRP; 09Z4J7; 1FLG.
CC InterPro: IPR001479; Bac_POQ.
CC InterPro: IPR002372; Bac_POQ_repeat.

```

```

DR InterPro: IPR000345; CytC_heme_bind.
DR Pfam: PF01011; Bacterial_POQ_6.
DR PROSITE: PS00363; BACTERIAL_POQ_1; 1.
DR PROSITE: PS00364; BACTERIAL_POQ_2; 1.
DR PROSITE: PS00190; CYTOCHROME_C_1.
KM Oxidoreductase; POQ; Heme; Periplasmic; Membrane; Signal.
FT SIGNAL 1 35
FT CHAIN 36 739
FT BINDING 651 651 HEME (COVALENT) (BY SIMILARITY).
FT BINDING 654 654 HEME (COVALENT) (BY SIMILARITY).
FT METAL 655 655 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
SQ SEQUENCE 739 AA; 80944 MW; E681BB237ACB91F4 CRC64;

Query Match 16.6%; Score 338; DB 1; Length 739;
Best Local Similarity 27.2%; Pred. No. 1,2e-18;
Matches 108; Conservative 54; Mismatches 185; Indels 50; Gaps 10;

OY 2 VANGVIAGSTCOYSPGCG--EYSGHDSATGEGELMRYFIIPRAGEGDETWGND----- 53
Db 204 IAKGRVLIIGN--GGSEFGARGFVYAFDAETGKVDMPNPNKNEPDHTASDVLMNKA 261
OY 54 YEARNMTGA-----WGQITYPVTNLVHYGSTAVGAPASTQCTGCTIGTINTR 103
Db 262 YQWSPFGAMTRQGGGCTWDSIYDPAVLVLGVNGSPWNYKRSSEKGNLFLGSI 321
OY 104 FAVRPDTGELVWRHQTLPNDNDQECTEFEMVNTVDVOPSTEMEGLQINPNATGERRV 163
Db 322 VALKPEYGEVYWHQETPRMDQWDTSVQIMTLDPINGET-----RHV 365
OY 164 LTGVPCKTGTMQFQDAETGPEFLWARDNTYQNMIESID-ENGIVYNEDAILKELDEYDV 222
Db 366 IVHAP-KNGFFYIIDAKTGEFISGKNVYVMAAGLDPKTGRPIYNDALTYTLGKEMYG 424
OY 223 CPTEFLGRDMPASALNDPSGIYPIPLNVCCYDMAVADQFT-SMDYNTS---NVTLP 278
Db 425 ITPDGLGHNFAAMAFSKTLVYIPAOQVPLTNOVGFTPHPDSSNNLGLDMKVGIPD 484
OY 279 G-----KMDIGRIDAIDISTGRTLSYERAANSPVLSYGGVLFNGTDRYFRAL 330
Db 485 SPEAKQAFVYKDLGWIYAWMPQKQAEAWRYVDHGMWNGILATGDLFLFGLANGEPHAY 544
OY 331 SQETGETIMQTRLATYASGQALSYEVDGMQYVIA 367
Db 545 DATNGSDLFHFAADSGIIPVYLYANGKQYVAVEVG 581

RESULT 6
DHET_ACEAC STANDARD; PRT; 742 AA.
AC P18278;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Alcohol dehydrogenase [acceptor] precursor (EC 1.1.99.8).
GN ADHA OR ADH1.
OS Acetobacter aceti.
OC Bacteria; Proteobacteria; alpha subdivision; Acetobacteraceae.
OX NCBI_TaxID=435;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 36-50.
RC MEDLINE=89255070; PubMed=2722742;
RA Inoue T., Sunagawa M., Mori A., Inai C., Fukuda M., Takagi M.,
RA Yano K.;
RT "Cloning and sequencing of the gene encoding the 72-kilodalton
RT dehydrogenase subunit of alcohol dehydrogenase from Acetobacter
RT aceti."
RL J. Bacteriol. 171:3115-3122(1989).
RN [2]
RP 3D-STRUCTURE MODELING.
RC MEDLINE=9528964; PubMed=7772016;
RA Cozler G.E., Giles I.G., Anthony C.;

```

RT "the structure of the quinoprotein alcohol dehydrogenase of
RT Acetobacter aceti modeled on that of methanol dehydrogenase from
RT Methylobacterium extorquens".
RL Biochem. J. 308:375-379(1995).
CC -1- CATALYTIC ACTIVITY: A primary alcohol + acceptor = an aldehyde +
CC reduced acceptor.
CC -1- COFACTOR: POQ AND HEME.
CC -1- SUBUNIT: TRIMER OF NON IDENTICAL CHAINS (DEHYDROGENASE,
CC CYTOCHROME, AND TWO SMALLER UNKNOWN SUBUNITS).
CC -1- SUBCELLULAR LOCATION: Periplasmic.
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POQ DEHYDROGENASE FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: D90004; BA14058.1; -
DR PIR: J50326; J50326.
DR HSP: Q924J7; IFUG.
DR InterPro: IPR001479; Bac_POQ.
DR InterPro: IPR002372; Bac_POQ_repeat.
DR InterPro: IPR000345; Cytc_heme_bind.
DR Pfam: PF01011; Bacterial_POQ_6.
DR PROSITE: PS00363; BACTERIAL_POQ_1; 1.
DR PROSITE: PS00364; BACTERIAL_POQ_2; 1.
DR PROSITE: PS00190; CYTOCHROME_C; 1.
DR Oxidoreductase; POQ; Heme; Periplasmic; Membrane; Signal.
FT SIGNAL 1 35
FT CHAIN 36 742 ALCOHOL DEHYDROGENASE [ACCEPTOR].
FT AC_SITE 343 343 BASE (POTENTIAL).
FT BINDING 649 649 HEME (COVALENT) (BY SIMILARITY).
FT BINDING 652 652 HEME (COVALENT) (BY SIMILARITY).
FT METAL 653 653 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
FT SEQUENCE 742 AA; 81521 MW; 96C9268DAB825A CnC64;
SO
Query Match 16.6%; Score 338; DB 1; Length 742;
Best Local Similarity 26.2%; Pred. No. 1.2e-18;
Matches 114; Conservative 67; Mismatches 152; Indels 102; Gaps 18;
OY 2 VANGIVAGSTQVSPFC--FVSGHDSATGEELMRNYFIPRAGECD----- 47
DB 203 VAKGIVLIGN--GGEFAPARGVSAFADETGLKMRFTVPNNKNEPDHAASDNILMKA 260
OY 48 -ETWGNDEYARMT-----GANGQITVPVNLVHYGSTAVGPASETORTGCTGLY--- 98
DB 261 YKTWGP--KGAMVRQGGGGTVDLSVYDPSDLIY---LAVG-----NGSPNNKYRSE 309
OY 99 ---GTN-----TRFVAVPDGGEIYMRHOTLPBDMWQECFEEMVAVVDVOPSTEMGLQSI 152
DB 310 GIGSLVLEIGSYALKEPGEIYMHFOATPMQMDYTSVOQITLTPVK----- 358
OY 153 NPNATGE-RVLTNGVPCKTGTMQFADTGEELIWARDNTNYOMTESID-----EN 202
DB 359 -----GEMRHVIVAP--KNGFEYVLDKATGTEFLSGKNVYVNMANGDPLGRPYND 411
OY 203 GIVTVNEADAIKELDEVYDVCPTFLGGRDWPALNPDGCIYFIPLNNYCI----- 253
DB 412 GLYTLNG-----KFWYGI--PGPLGHNFMAMAYSPKTHLVIPAHOIPEFYKNVGGF 463
OY 254 -----DMNAVDOEFTSMDYNTNSVTKLPKCKMIGRIDIDISTGTLMSVRAAANYSP 309
DB 464 KPHASWNVGLDMTKNGLDIPE-ATAYIKIDHGLWLMWDVYKMEYVAKIDHKGWNGG 522
OY 310 VLTGGGVLENGCTRRYFALSOETGETLMOQTRLATVASGAISYEVDQMYVA----- 363
DB 523 ILATGSDLLFOGLANGEFHAYDATNGSDLYKFDASGIIAPMTYSVNGKQYVAIVEVGW 582
OY 364 ----TAGGVSYSYGSG 374

DB 583 GIYPLSMGVGRTSG 597
11: 111 11
RESULT 7
ID DHM1_METME STANDARD: PRT; 571 AA.
AC P38539;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Methanol dehydrogenase subunit 1 (EC 1.1.99.8) (MDH large alpha
DE subunit) (MEDH).
OS Methylophilus methylotrophus (Bacterium M3A1).
OC Bacteria; Proteobacteria; beta subdivision; Methylophilus group;
OC Methylophilus.
OX NCBI_TaxID=17;
RN [1]
RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS).
RX MEDLINE=94059969; PubMed=8241148;
RA White S., Boyd G., Mathews F.S., Xia Z.-X., Dai W.-W., Zhang Y.-F.,
RA Davidson V.L.;
RT "The active site structure of the calcium-containing quinoprotein
RT methanol dehydrogenase".
RL Biochemistry 32:12955-12958(1993).
RN [2]
RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS).
RX MEDLINE=93054513; PubMed=131050;
RA Xia Z.-X., Dai W.-W., Xiong J.-P., Hao Z.-P., Davidson V.L., White S.,
RA Mathews F.S.;
RT "The three-dimensional structures of methanol dehydrogenase from two
RT methylotrophic bacteria at 2.6-A resolution".
RL J. Biol. Chem. 267:22289-22297(1992).
CC -1- CATALYTIC ACTIVITY: A primary alcohol + acceptor = an aldehyde +
CC reduced acceptor.
CC -1- COFACTOR: TWO MOLECULES OF POQ AND TWO MOLECULES OF CALCIUM
CC PER TRIMER.
CC OF A LARGE ALPHA AND A SMALL BETA SUBUNIT.
CC -1- SUBCELLULAR LOCATION: PERIPLASMIC, BUT ASSOCIATED WITH INNER
CC MEMBRANE.
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POQ DEHYDROGENASE FAMILY.
CC PDB: 4AAH; 08-DEC-96.
DR InterPro: IPR001479; Bac_POQ.
DR InterPro: IPR002372; Bac_POQ_repeat.
DR Pfam: PF01011; Bacterial_POQ_7.
DR PROSITE: PS00363; BACTERIAL_POQ_1; FALSE_NEG.
DR PROSITE: PS00364; BACTERIAL_POQ_2; 1.
KM Oxidoreductase; POQ; Methanol utilization; Periplasmic; 3D-structure;
KM Calcium.
FT DISULFID 103 104
FT DISULFID 379 408
FT AC_SITE 297 297 BASE (POTENTIAL).
FT HELIX 2 9
FT TURN 11 12
FT STRAND 14 14
FT TURN 17 18
FT TURN 21 22
FT STRAND 26 27
FT TURN 34 36
FT TURN 37 39
FT HELIX 41 47
FT STRAND 49 61
FT STRAND 62 63
FT STRAND 64 68
FT TURN 71 73
FT TURN 75 79
FT STRAND 80 81
FT TURN 83 84
FT STRAND 86 90
FT TURN 96 101
FT HELIX 103 104
FT TURN 103 104
FT STRAND 112 114

FT TURN 115 116
 FT STRAND 117 121
 FT TURN 123 124
 FT STRAND 126 131
 FT TURN 132 134
 FT STRAND 137 142
 FT HELIX 146 148
 FT TURN 149 149
 FT STRAND 151 151
 FT STRAND 157 159
 FT TURN 160 161
 FT STRAND 162 166
 FT STRAND 168 168
 FT HELIX 170 172
 FT TURN 173 173
 FT STRAND 177 182
 FT TURN 183 185
 FT STRAND 188 193
 FT TURN 198 203
 FT STRAND 203 203
 FT TURN 205 210
 FT TURN 212 213
 FT HELIX 223 223
 FT TURN 226 231
 FT STRAND 241 243
 FT TURN 244 247
 FT STRAND 248 252
 FT HELIX 261 263
 FT TURN 269 270
 FT STRAND 273 278
 FT TURN 279 281
 FT STRAND 284 289
 FT STRAND 305 311
 FT TURN 312 313
 FT STRAND 314 322
 FT TURN 324 325
 FT STRAND 327 332
 FT TURN 333 335
 FT STRAND 336 343
 FT TURN 346 347
 FT STRAND 351 355
 FT TURN 356 359
 FT STRAND 360 363
 FT HELIX 365 367
 FT TURN 370 370
 FT STRAND 371 372
 FT STRAND 373 373
 FT STRAND 375 378
 FT TURN 382 383
 FT STRAND 390 390
 FT STRAND 392 394
 FT TURN 395 398
 FT STRAND 399 405
 FT STRAND 407 413
 FT TURN 420 421
 FT STRAND 427 433
 FT TURN 435 436
 FT TURN 438 439
 FT STRAND 444 450
 FT TURN 451 454
 FT STRAND 455 462
 FT STRAND 469 469
 FT STRAND 471 473
 FT TURN 474 476
 FT STRAND 477 481
 FT TURN 483 484
 FT STRAND 486 491
 FT TURN 492 495
 FT STRAND 496 502
 FT STRAND 512 516
 FT TURN 517 518
 FT STRAND 519 526

FT TURN 530 533
 FT HELIX 534 538
 FT TURN 539 539
 FT TURN 543 545
 FT HELIX 546 552
 FT TURN 553 555
 FT HELIX 556 558
 FT TURN 559 559
 FT STRAND 565 570
 SQ SEQUENCE 571 AA: 62449 MM: 0BE94EA5AD2AB1E1 CRC64;

Query Match 16.4%; Score 332.5; DB 1; Length 571;
 Best local similarity 25.1%; Pred No. 2, 3e-18;
 Matches 111; Conservative 65; Mismatches 161; Indels 105; Gaps 18;

QY 3 ANGVIVA-----GSTCQYSPF-----GC-----FVSGHD 26
 DB 123 ANGHLLALDAKTGKINWEVEVCDPKVGSFLTQAPFVAKDTVLGMSGAGELGVAGVNAFD 182
 QY 27 SATGEELMR-----NYFIPRAGE--EGDETGNQDYEARMTGA---WQ 65
 DB 183 LKTGELKWRAPATGSDSVRLAKDFNSANDHYGQFGLGKTWEGD---AMKIGGTMGW 239
 QY 66 ITYDPYTNLVHVGSTAVGASETQGTGPG-----TLVGNTRFAVRPPTGELVHRHOTL 120
 DB 240 YADPKLNFYIOSSGNPAPWNETMR--PGDNKKTMTIWGRDL-----DTGAKKQYQKT 291
 QY 121 PRDNWDECTFEMKMTNVDP--STEMEGLQS--INPNAATGERRVLTGVPCKTGMQPD 178
 DB 292 PHDEMDFAGVNMVLTLD---QPVNAKTFLLSHIDRN-----GLYTLN 332
 QY 179 AETGEFLMARDTN-YQNMIESIDENGIYVNEALIKELDVE-YDVCPTFLGRDMPSPA 236
 DB 333 RENGNIIVAEKYPDAVAVRKKVDLKTGTPVRODEPFRTRMDHGTNCPSPAMGHNGVDS 392
 QY 237 LNPDSGIVFPLNNWCYDMAVAVDDEFTSMQVNTSNTKLP---PGKMDIGRIDADIS 292
 DB 393 YDPESRTLVAGLNHICMDNEPFLPYRAGQFVGATLAMPNGNNGTKKEMQIRAFDLT 452
 QY 293 TGRITMSVERAANYSPIVSTGGVLFNGGTRFALSOEGETIMQTRLATVASQAI 352
 DB 453 TGRAKWTWKEKPAWGGTLTYTGGLVWATLDGYLKALDNKQKELMKNKMSGSGISPM 512
 QY 353 SYEVDGMOYVAIAGGVSYSG 374
 DB 513 TYSFKGQYI-----GSMGVG 529

RESULT 8
 DHET_ACEPO
 ID DHET_ACEPO STANDARD: PRT: 738 AA.
 AC P28036:
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Alcohol dehydrogenase [acceptor] precursor (EC 1.1.99.8).
 GN ADHA.
 OS Acetobacter polyoxogenes.
 OC Bacteria; Proteobacteria; alpha subdivision; Acetobacteraceae;
 OC Acetobacter.
 OX NCBI_TaxID=439;
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC STRAIN=NB11028;
 RX MEDLINE=91159482; PubMed=2001402;
 RA Tamaki T., Fukaya M., Takemura H., Tayama K., Okumura H.,
 RA Kawamura Y., Nishiyama M., Horinouchi S., Beppu T.;
 RT "Cloning and sequencing of the gene cluster encoding two subunits of
 membrane-bound alcohol dehydrogenase from Acetobacter polyoxogenes.";
 RL Biochim. Biophys. Acta 1088:292-300(1991).
 CC -1- FUNCTION: DEHYDRATION OF PRIMARY ALCOHOLS (EXCEPTION: METHANOL).
 CC -1- CATALYTIC ACTIVITY: A primary alcohol + acceptor -> an aldehyde +

DR PROSITE: PS00364; BACTERIAL_POQ_2; FALSE_NEG.
 KW Oxidoreductase; POQ; Periplasmic; Signal; Calcium; 3D-structure;
 KM Complete proteome.
 FT SIGNAL 1 34
 FT CHAIN 35 623 QUNOBIOTIN ETHANOL DEHYDROGENASE.
 FT DISULFID 139 140
 SQ SEQUENCE 623 AA; 68123 MW; 32DE5DF20B291D6 CRC64;

Query Match 15.2%; Score 308.5; DB 1; Length 623;
 Best Local Similarity 22.9%; Pred. No. 1.9e-16;
 Matches 88; Conservative 68; Mismatches 175; Indels 53; Gaps 10;

QY 26 DSAAGELMKNYFPR-----AGEEDETWMDYEA-----RWMTGA--WG 64
 DB 224 DPTDGETMRRPVEGGMGLNGKSDSTVGDYKAPSPDRNSPTGKVSWSHGGAPOQ 283
 QY 65 QITVDPTNLVHYGSTVAVPASETORGTGG-----TLTGTWTRAVRPDGEIYWRHQ 118
 DB 284 SASFDAETNLIYAGAGPMPWMTWARTAKGNPHDSDLT-ISGQGVDPSSGEVAKWFTQ 342
 QY 119 TLRDNDQDCTEEMAVTVNDVOPSTEMEGLOSINPN-----AATGERRVLGVPCKTGT 173
 DB 343 HTPNDAWDESGNNELVDFYKAKDKIKVATAHADNNGEYVYDRSNGKLQNAFPVDNI 402
 QY 174 MW--OFDAETGEFLMARDTNYOMIESIDENGIVTVNEDAIKELDVEYDVCPTFLIGRD 231
 DB 403 TMSHIDDKTG-----RPVEREQRPPLDPCG-----QKHGAVEVSPPLFGGKN 447
 QY 232 WPSAALNPDGSIYFIPLNVCYDMAVDOFTSMVYNTSNYTKLPPGKDMIRDAIDI 291
 DB 448 WNNMAYSDPTGLFVYFANHMKEDYWEVSYTKGSAYLGMGRFKRMKYDHDHGLRAMDP 507
 QY 292 STGRTLSVERAANSPVLSITGGVLENGCDRIYFRALSOETGELTMOIRATVYASGQA 351
 DB 508 VSKVIAWEKHEHLPLMAVGLATAGNLVFTGTDGYFKAPDASGKELMFKTGSGIVSPP 567
 QY 352 ISYVDGMQVYAIAGGVSYSGL 375
 DB 568 ITWEDGBOYLGYT---VGTGAV 588

RESULT 10
 DHET_GLUOX STANDARD; PRT; 757 AA.
 AC 005542;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Alcohol dehydrogenase [acceptor] precursor (EC 1.1.99.8) (G3-ADH subunit 1).
 GN ADHA.
 OS Gluconobacter oxydans (Gluconobacter suboxydans).
 OC Bacteria; Proteobacteria; alpha subdivision; Acetobacteraceae;
 OC Gluconobacter.
 OX NCBI_TaxID=442;
 RP SEQUENCE FROM N.A. AND SEQUENCE OF 36-50.
 RC STRAIN-IFO 12528;
 RX MEDLINE-97208225; PubMed-9055427;
 RA Kondo K., Horinouchi S.;
 RT "Characterization of the genes encoding the three-component membrane-bound alcohol dehydrogenase from Gluconobacter suboxydans and their expression in Acetobacter pasteurianus";
 RL Appl. Environ. Microbiol. 63:1131-1138(1997).
 CC -1- FUNCTION: DEHYDRATION OF PRIMARY ALCOHOLS (EXCEPTION: METHANOL) (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: A primary alcohol + acceptor = an aldehyde + reduced acceptor.
 CC -1- COFACTOR: POQ AND HEME.
 CC -1- SUBUNIT: HETEROTRIMER (DEHYDROGENASE, CYTOCHROME AND PROTEIN ADHS), THAT FORMS THE ALCOHOL DEHYDROGENASE COMPLEX.
 CC -1- SUBCELLULAR LOCATION: MEMBRANE-BOUND, FACING THE PERIPLASMIC

CC SPACE (POTENTIAL).
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POQ DEHYDROGENASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: D86375; BA019753.1; -.
 DR HSBP: Q9Z4U7; ILEG.
 DR InterPro: IPR001479; Bac_POQ.
 DR InterPro: IPR002372; Bac_POQ_repeat.
 DR InterPro: IPR000345; Cytochrome_bind.
 DR Pfam: PF01011; Bacterial_POQ_6.
 DR PROSITE: PS00363; BACTERIAL_POQ_1.
 DR PROSITE: PS00364; BACTERIAL_POQ_2; 1.
 DR PROSITE: PS00190; CYTOCHROME_C; 1.
 KW Oxidoreductase; POQ; Heme; Periplasmic; Membrane; Signal.
 FT SIGNAL 1 34
 FT CHAIN 35 757 ALCOHOL DEHYDROGENASE [ACCEPTOR].
 FT MOD_RES 35 35 PYRROLIDONE CARBOXYLIC ACID.
 FT ACT_SITE 342 342 BASE (POTENTIAL).
 FT BINDING 653 653 HEME (COVALENT) (BY SIMILARITY).
 FT BINDING 656 656 HEME (COVALENT) (BY SIMILARITY).
 FT METAL 657 657 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
 SQ SEQUENCE 757 AA; 82968 MW; 39B9F903B947581 CRC64;

Query Match 14.7%; Score 298.5; DB 1; Length 757;
 Best Local Similarity 23.2%; Pred. No. 1.5e-15;
 Matches 99; Conservative 68; Mismatches 180; Indels 79; Gaps 13;

QY 2 VANGIYAGSCQSPFGCC--FVSGHDSATGEELMRYFIIPRAGEED----- 47
 DB 202 IAKGKVLGN--GGAERFAGSFSAFDEFSKLDWRFYVFNPNKPKDGAASDILMSKA 259
 QY 48 -ETWGNDEYEARMTG---AMGQITVDPTNLVHYGSTVAVPASETORGTGGTLTGNTNR 103
 DB 260 YPTGKNGAMKQGGGGGVMDSLVYDPTDLYLVGVNGSPMWKFRSEKGNLFLGSI 319
 QY 104 FAVRPDGEIYWRHQITPLPRDNDQDCTEEMAVTVNDVOPSTEMEGLOSINPNAATGE-RR 162
 DB 320 VAINPDGKYVMEHQETPRMDEMDTVSQQITLIDMPV-----NEMERH 362
 QY 163 VLTGVPCRTGTMQFOFADTGEFLMARDTNYOMIESIDE-NGIVTVNEDAIKELDVEYD 221
 DB 363 VIVHAP-KNGEFTIIDAKTKGFTGKRYTEANNANGILDPTGPRNYVPDLTLTGKPMW 421
 QY 222 VCPFLGGRDWPSSAALNPDGSIYFIPLNV-----CYDMAVDOFTSMVY 268
 DB 422 GIPBELGHNFAAMAYSPKTKLVIPAOQIPLLYDQGGKFAYHDAMNGLDMNKIGLF 481
 QY 269 NTSNVTKLPKGMIGRIDAIDISTGRTL-WSEVERAANVS-----PYLSTGGSVL 318
 DB 482 DDNDDEHVAKKDF-----LKVLKGTVAWDEKAPAFETINHKPKPMNGGLATAGNYI 535
 QY 319 FNGSTDRYFRALSOETGELTMOIRATVYASGQAISYVDGMQVYI-----AGGG 368
 DB 536 FQGLANEFHAYATNGNDLISFPASAIITAPVYTTANGKOYVAVEVGGIIPPLYGG 595
 QY 369 VSYGSG 374
 DB 596 VARTSG 601

RESULT 11
 DHGA_ACIQA STANDARD; PRT; 801 AA.
 AC P05465;
 ID DHGA_ACIQA
 DT 01-NOV-1988 (Rel. 09, Created)

DT 01-NOV-1988 (Rel. 09, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Glucose dehydrogenase-A [pyrroloquinoline quinone] precursor
 DE (EC 1.1.99.17) (Quinoprotein glucose DH) (GDH-A).
 GN GDH.
 OS Acinetobacter calcoaceticus.
 OC Bacteria; Proteobacteria; gamma subdivision; Moraxellaceae;
 OC Acinetobacter.
 OC NCBI_TaxID=471.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=LMD 79.41;
 RX MEDLINE=88289368; PubMed=3393933;
 RA Cleton-Jansen A.-M., Goosen N., Odle G., van de Putte P.;
 RT "Nucleotide sequence of the gene coding for quinoprotein glucose
 dehydrogenase from Acinetobacter calcoaceticus.";
 RL Nucleic Acids Res. 16:6228-6228(1988).
 CC -1- CATALYTIC ACTIVITY: D-glucose + acceptor = D-glucono-1,5-lactone +
 reduced acceptor.
 CC -1- COFACTOR: POQ.
 CC -1- SUBUNIT: MONOMER.
 CC -1- SUBCELLULAR LOCATION: Periplasmic.
 CC -1- MISCELLANEOUS: ACINETOBACTER CALCOACETICUS CONTAINS TWO DIFFERENT
 POQ DEPENDENT ENZYMES WITH GDH ACTIVITY. GDH-A PREFERENCES
 2-DEOXYGLUCOSE AS SUBSTRATE, THE SPECIFIC SUBSTRATES FOR GDH-B ARE
 DISACCHARIDES.
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POQ DEHYDROGENASE FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation
 at the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@sib-sib.ch).
 CC
 CC EMBL: X07235; CA30222.1; -.
 DR PIR: S00943; S00943.
 DR InterPro: IPR001479; Bac_POQ.
 DR Pfam: PF01011; Bacterial_POQ_7.
 DR PROSITE: PS00363; BACTERIAL_POQ_1; 1.
 DR PROSITE: PS00364; BACTERIAL_POQ_2; 1.
 KW Oxidoreductase; POQ; Transmembrane; Periplasmic; Signal.
 FT SIGNAL 1 33
 FT CHAIN 34 801
 FT TRANSMEM 39 55
 FT TRANSMEM 59 79
 FT TRANSMEM 94 108
 FT TRANSMEM 119 138
 FT ACT_SITE 471 471
 FT SEQUENCE 801 AA; 86956 MW; 2F4EA222EB5450D8 CRC64;
 SQ
 Query Match 12.5%; Score 253; DB 1; Length 801;
 Best Local Similarity 23.0%; Pred. NO. 5,4e-12;
 Matches 113; Conservative 62; Mismatches 148; Indels 168; Gaps 25;
 QY 1 IVANGVIAGS-TCOYS---PFGCFVSGHSATGSEELMRNFIPRAGE-----EGDETW 50
 DB 346 VTGSTVIVAGSVTNYNSKKEPSG-VIRQYDVNTKILM--VFDTGADPNAMPGEITFEV 402
 QY 51 GNDYEARMGTAGCOITVDPTNLVHGSTAVGPASSETORCTPGCTLYG-----T 100
 DB 403 HNS-----PANMAPLAVDALDLIV-VYPTGV-----GTP--DLWGGDRTELKERYA 445
 QY 101 NTRFAVRPDTGEIYWRHOTLPDRMDDECTETEMAVTNDVPSSTEMEGLOSIN--NAAT 158
 DB 446 NSMLAINASTKLVNFOPTHDLMDVPSQPSLADIKNAGQVPAIYVLTFTGNAFV 505
 QY 159 GERRVLTVGP-----CTG--TMMQFDETFGEF-----LWA-- 187

DB 506 LDRR--NCQPIVPTVEKRPOTVAKRGQPTKGEFYSKTPQPSFDLAPQDKLTDKDMWGAT 563
 QY 188 -----RDNTYQNMIESIDENGIVYNEDAILKELDEVDCPTFLGGDMPSAA 236
 DB 564 MLDQMLCRVSEFKRLNLYDITPPSENGTL-----VPPGNLGVSEWGMS 607
 QY 237 LNPDS-----GLYFI-----PLNNVCYDMAVADQERTS 264
 DB 608 VNPDRQVAVMNPDIGLPFYSRLIPADPNRAQAKAGTEQGVPMYGVFY-----G 657
 QY 265 MDVYNTSNVTKLPCKDMIGRIDAIDISTGRTLW-----SVERAAN 306
 DB 658 VEISAFSLPLGPKQPMAGVAGVDLKTHEVWKKRIGTIRDSLPLNFQLPAAKIGVPG 717
 QY 307 YSPVLSTGGVLENGGT--DRYFRALISOETGETTMOGRLATVASQA--ISTEVDMQYVA 363
 DB 718 LGGSTAGNVAVFVATQDNYLAFNVYNGRKLWEARLP--AGQADPMTEYINGKQYV 775
 QY 364 I-AGGVSYSYS 373
 DB 776 IMAGHGSFRT 786
 RESULT 12
 DHG_ECOLI STANDARD: PRT; 796 AA.
 ID DHG_ECOLI
 AC P15877;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Glucose dehydrogenase [pyrroloquinoline-quinone] (EC 1.1.99.17).
 GN GCD OR B0124.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.
 OX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12;
 RX MEDLINE=91035240; PubMed=2228962;
 RA Cleton-Jansen A.-M., Goosen N., Fayet O., van de Putte P.;
 RT "Cloning, mapping, and sequencing of the gene encoding Escherichia
 coli quinoprotein glucose dehydrogenase.";
 RL J. Bacteriol. 172:6308-6315(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / W3110;
 RX MEDLINE=93123180; PubMed=8419307;
 RA Yamada M., Asaoka S., Sailer M.H. Jr., Yamada Y.;
 RT "Characterization of the gcd gene from Escherichia coli K-12 W3110
 and regulation of its expression.";
 RL J. Bacteriol. 175:568-571(1993).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / W3110;
 RX MEDLINE=94261430; PubMed=8202364;
 RA Fujita N., Mori H., Yura T., Ishihama A.;
 RT "Systematic sequencing of the Escherichia coli genome: analysis of
 the 2.4-4.1 min (110,917-193,643 bp) region.";
 RL Nucleic Acids Res. 22:1637-1639(1994).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / MC1655;
 RX MEDLINE=97426617; PubMed=9278503;
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of Escherichia coli K-12.";
 RL Science 277:1453-1474(1997).
 RN [5]
 RP TOPOLOGY.
 RX MEDLINE=93286127; PubMed=8509415;

RA Yamada M., Sumi K., Matsushita K., Adachi O., Yamada Y.;
 RT "Topological analysis of quinoprotein glucose dehydrogenase in
 RT Escherichia coli and its ubiquinone-binding site.";
 RN J. Biol. Chem. 268:12812-12817(1993).
 RL [6]
 RP 3D-STRUCTURE MODELING.
 RX MEDLINE:96128046; Pubmed-8554505;
 RA Cozler G.E., Anthony C.;
 RT "Structure of the quinoprotein glucose dehydrogenase of Escherichia
 RT coli modelled on that of methanol dehydrogenase from Methylobacterium
 RT extorquens";
 RL Biochem. J. 312:679-685(1995).
 CC -1- FUNCTION: GDH IS PROBABLY INVOLVED IN ENERGY CONSERVATION RATHER
 CC THAN IN SUGAR METABOLISM.
 CC CATALYTIC ACTIVITY: D-glucose + acceptor = D-glucono-1,5-lactone +
 CC reduced acceptor.
 CC -1- COFACTOR: POQ.
 CC -1- SUBUNIT: MONOMER.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane.
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POQ DEHYDROGENASE FAMILY.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-slb.ch/announce/>
 CC or send an email to license@isb-slb.ch).
 CC -----
 DR EMBL: X51323; CAA35706.1; -
 DR EMBL: D12651; BAA02174.1; -
 DR EMBL: D26562; CAB20298.1; -
 DR EMBL: AE000122; AAC73235.1; -
 DR PIR: JY0107; JY0107.
 DR HSSP: P38539; 4AAH.
 DR EcoGene: EG10369; gcd.
 DR InterPro: IPR001479; Bac_POQ.
 DR InterPro: IPR002372; Bac_POQ_repeat.
 DR Pfam: PF01011; Bacterial_POQ_7.
 DR PROSITE: PS00363; BACTERIAL_POQ_1; 1.
 DR PROSITE: PS00364; BACTERIAL_POQ_2; 1.
 DR Oxioreductase; POQ; Transmembrane; Inner membrane; Periplasmic;
 KW Complete proteome.
 FT DOMAIN 1 10 CYTOPLASMIC (PROBABLE).
 FT TRANSMEM 38 37 PROBABLE.
 FT DOMAIN 38 40 PERIPLASMIC (PROBABLE).
 FT TRANSMEM 41 58 PROBABLE.
 FT DOMAIN 59 62 CYTOPLASMIC (PROBABLE).
 FT TRANSMEM 63 81 PROBABLE.
 FT DOMAIN 82 95 PERIPLASMIC (PROBABLE).
 FT TRANSMEM 96 110 PROBABLE.
 FT DOMAIN 111 118 CYTOPLASMIC (PROBABLE).
 FT TRANSMEM 119 141 PROBABLE.
 FT DOMAIN 142 796 PERIPLASMIC (PROBABLE).
 FT ACT_SITE 466 466 BASE (PROBABLE).
 FT CONFLICT 59 59 R -> L (IN REF. 1).
 FT CONFLICT 149 156 TUSADATP -> HIKRRCHT (IN REF. 1 AND 2).
 FT CONFLICT 193 193 N -> K (IN REF. 1 AND 2).
 FT CONFLICT 666 666 Q -> H (IN REF. 2).
 SQ SEQUENCE 796 AA; 86747 MW; D9EDC705A12894E5 CMC64;

Query Match 11.4%; Score 231.5; DB 1; Length 796;
 Best Local Similarity 23.4%; Pred. No. 2.5e-10;
 Matches 113; Conservative 53; Mismatches 154; Indels 163; Gaps 23;

OY 6 VIVAGS-TCOXS--PFGFVSGHSATSEELMRNFIRPRA-----GEEDDFWGDNDYE 55
 DB 346 IYMASVTDNSTRRTSCVIRKFDVNTGELLMA--FDGADAPNAIPDEDTFFNS--- 400
 OY 56 ARMTGANGQITDPVFNLMVHYGSTAVGPASETORGPGTLYGTNFFAVRPDTGETIYW 115
 DB 401 ----PNSWAPAPAYDAKDLVLLPVGVTTPDIMGWNRPEQGRVASSI-LALNATTKGLAW 455

OY 116 RHQITLPRDWDQCEFFEMAVTVNDVQPSFEMEGLOSINPNATGRRVLTGVPCKTGPMW 175
 DB 456 STQYVHHDDMDND-----LPAQPT-----LADIVYN---GQKVVIYAPAKTGNIFF 498
 OY 176 QPDAETGEFL-----WARDNTQNMIESIDENGI-----VYVNEDAI 212
 DB 499 VLDNRNGELVYAPERKPVYQGAARADYVTPPOFSELSFRPRKDLGADMGATVFDGLV 558
 OY 213 LKEL--DVEYD-----VCPPLGRDWSAALNP----- 239
 DB 559 CVMFQHMRYEGIFTPPSBOGLVFPNGIMHEMGISVDNPREVALIANPALPFSKLI 618
 OY 240 -----DSGI-----YFIPLNVCYDMAVDOEFTSMDYVNTSNV 273
 DB 619 PRGPGNPMPQPDANGCTGESGIPQYGVPPVGTLTN-----PFLS-----P 659
 OY 274 TKLPKQDMIGRIDAIDISTGRTLSVERAANYS-----PV-----LSTG 314
 DB 660 FELPKQKAPAMGYISALDLKTNEVYWKRIQIPQDSMPMPVPPVFNMGMPNLGGPISTA 719
 OY 315 GGVLENGGT-DRIYFALSQETGELIMQTRATVASGQA--ISYEVDQMUYAI-AGGGSV 370
 DB 720 GNVLEIATADNYLRAVNMKSGEKIMQGRLP--AGGQATPMYEVNGKQYVIVASGGHS 777
 OY 371 YGS 373
 DB 778 FGT 780

RESULT 13
 QUITA_ACICA STANDARD; PRT; 809 AA.
 ID QUITA_ACICA
 AC 059086;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Quinate/shikimate dehydrogenase [Pyroloquinoline-quinone]
 DE (EC 1.1.99.25) (NAD(P)-independent quinate dehydrogenase).
 GN QUITA.
 OS Acinetobacter calcoaceticus.
 CC Bacteria; Proteobacteria; gamma subdivision; Moraxellaceae;
 CC Acinetobacter.
 OX NCBI_TaxID=471;
 RN [1]
 RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
 RC STRAIN-BD413 / ADP1.
 RX MEDLINE-9505936; Pubmed-8002591;
 RA Elsemore D.A., Ornstion L.N.;
 RT "The pca-pob supraperoic cluster of Acinetobacter calcoaceticus
 RT contains quia, the structural gene for quinate-shikimate
 RT dehydrogenase.";
 RL J. Bacteriol. 176:7659-7666(1994).
 RN [2]
 RN SEQUENCE OF 1-20 FROM N.A.
 RC STRAIN-BD413 / ADP1;
 RX MEDLINE-96011389; Pubmed-7592351;
 RA Elsemore D.A., Ornstion L.N.;
 RT "Unusual ancestry of dehydratases associated with quinate catabolism
 RT in Acinetobacter calcoaceticus.";
 RL J. Bacteriol. 177:5971-5978(1995).
 CC -1- FUNCTION: CAN ACT EITHER ON QUINATE OR ON SHIKIMATE.
 CC -1- CATALYTIC ACTIVITY: Quinate + pyroloquinoline-quinone = 5-
 CC dehydroquinate + reduced pyroloquinoline-quinone.
 CC -1- CATALYTIC ACTIVITY: SHIKIMATE + PYROLOQUINOLINE-QUINONE = 3-
 CC DEHYDROSHIKIMATE + REDUCED PYROLOQUINOLINE-QUINONE.
 CC -1- COFACTOR: POQ.
 CC -1- PATHWAY: QUINIC ACID CATABOLIC PATHWAY; FIRST STEP. THIS PATHWAY
 CC ALLOWS GROWTH OF BACTERIA WITH QUINATE BY ITS CONVERSION TO
 CC PROTOCATECHUATE AND SUBSEQUENT METABOLISM BY THE BETA-KETOADIPATE
 CC PATHWAY.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (potential).
 CC -1- INDUCTION: BY PROTOCATECHUATE.

CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POQ DEHYDROGENASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: L05770; AAC37161.1; -
 CC InterPro: IPR001479; Bac_PQO.
 CC Pfam: PF01011; Bacterial_PQO_7.
 CC PROSITE: PS00363; BACTERIAL_POQ_1; FALSE_NEG.
 CC PROSITE: PS00364; BACTERIAL_POQ_2; FALSE_NEG.
 CC Oxidoreductase; POQ; Quinate metabolism; Transmembrane.
 CC TRANSMEM 14 34 POTENTIAL.
 CC TRANSMEM 41 61 POTENTIAL.
 CC TRANSMEM 68 88 POTENTIAL.
 CC TRANSMEM 90 110 POTENTIAL.
 CC TRANSMEM 127 147 POTENTIAL.
 CC SEQUENCE 809 AA; 88196 MW; 71F67CEBBA62BFCB CRC64;
 SQ
 Query Match 10.1%; Score 204.5; DB 1; Length 809;
 Best Local Similarity 22.9%; Pred. No. 3,1e-08;
 Matches 105; Conservative 54; Mismatches 187; Indels 113; Gaps 17;
 QY 2 VANGVIVAGSTCYSPF-----GCFVSGHDSATGELMKNYFPRAGGEGDETWGNDYAR 57
 DB 354 IAGTIVGSRADNVADMPGVRADYVITGKLRMA--EDPRNDPNYLYKPEIKR 411
 QY 58 WMTGAMGOITYPVNLVH--VGTAVGPASRGTGCTGTGNTFFAVAPDGEITW 115
 DB 412 SSTNSMAASIDPOMNTVFLPMGSSVDYWGGR--TAADHYNTSV--LALDATTGKRW 468
 QY 116 RHQTLPRDNDDECTFEMNTVDVQPTSTEMGLQSIENFNAATGERVLTGPKCTGTW 175
 DB 469 VNTVAVNDLMD-----FDLPMPQS-----LVDFPKDGTTPAVVIGT--KSGQFY 512
 QY 176 QPDAETGEFL-----WADPTNYONMIESIDENIVVNE-----DAI 212
 DB 513 VLDRTVKGKPLRVIEPIKADIPGEYSKTPRSVEMQIQNQLKESDMWGAPPPDL 572
 QY 213 LKELD--VEVD-----VCPFLGGRDPSALNDPSIYF-----245
 DB 573 MCRINKSKRYDGLTAPGTDSLSFPGSLGGMNGSIAPDTHRYMVFNDRLGLMTQL 632
 QY 246 -----IPLN--NVCYDMNAVDEFTSMQVNTSVTKL--PPGKDMIGRIDAI 291
 DB 633 IKQTPEDIKIQANGKBEKVTGMGAVPMKGTPIKVNKNRPMFSAIGIPCOKPPFGTMTADM 692
 QY 292 STGRITLMSVERAANS-----PVLSTGGGVLFGNGDRYERALS 331
 DB 693 KTRQVAMQVPLGTIDTQDTPMGIKMGLKADIGAPMTIGGPRATGGGLVFATQDYILRAF 752
 QY 332 QETGETIMOTRLATVASGOAISY--EVDGMOYVAIAGG 368
 DB 753 SSKGELMKARLPVSGGTPMSPKTKQYVVSAGG 791
 RESULT 14
 DHG_GLUOX STANDARD: PRT; 808 AA.
 AC P27175;
 DT 01-AUG-1992 (rel. 23, Created)
 DT 01-AUG-1992 (rel. 23, Last sequence update)
 DT 16-OCT-2001 (rel. 40, Last annotation update)
 DE Glucose dehydrogenase [pyrroloquinoline-quinone] precursor
 DE (EC 1.1.99.17).
 GN Gdh.
 OS Gluconobacter oxydans (Gluconobacter suboxydans).

OC Bacteria; Proteobacteria; alpha subdivision; Acetobacteraceae;
 OC Gluconobacter.
 OC NCBI_TaxID=442;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-92017653; PubMed-1833618;
 RA Cleton-Jansen A.-M., Dekker S., van de Putte P., Goosen N.;
 RT "A single amino acid substitution changes the substrate specificity
 RT of quinoxaline glucose dehydrogenase in Gluconobacter oxydans";
 RL Mol. Gen. Genet. 229:206-212(1991).
 RN [12]
 RP NEWSTON TO 213.
 RA Goosen N.;
 RL Submitted (FEB-1992) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: D-glucose + acceptor = D-glucono-1,5-lactone +
 CC reduced acceptor.
 CC -1- COFACTOR: POQ.
 CC -1- SUBCELLULAR LOCATION: OUTSIDE OF THE PERIPLASMIC MEMBRANE.
 CC -1- MISCELLANEOUS: THE P1 FORM CAN OXIDIZE ONLY D-GLUCOSE, WHILE THE
 CC P2 FORM CAN ALSO OXIDIZE MALTOSE. THE SEQUENCE OF P1 FORM IS SHOWN
 CC HERE.
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POQ DEHYDROGENASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: X62710; CAN44594.1; ALT_SEQ.
 CC PIR: S17716; QPKEX.
 CC InterPro: IPR001479; Bac_PQO.
 CC InterPro: IPR002372; Bac_PQO_repeat.
 CC Pfam: PF01011; Bacterial_PQO_7.
 CC PROSITE: PS00363; BACTERIAL_POQ_1; 1.
 CC PROSITE: PS00364; BACTERIAL_POQ_2; 1.
 CC Oxidoreductase; POQ; Transmembrane; Periplasmic; Signal.
 CC KW SIGNAL 1 33 POTENTIAL.
 CC FT CHAIN 34 808 GLUCOSE DEHYDROGENASE [PYRROLOQUINOLINE-
 CC QUINONE].
 CC FT TRANSMEM 35 54 POTENTIAL.
 CC FT TRANSMEM 59 76 POTENTIAL.
 CC FT TRANSMEM 94 108 POTENTIAL.
 CC FT TRANSMEM 123 138 POTENTIAL.
 CC FT ACT_SITE 470 470 BASE (POTENTIAL).
 CC FT VARIANT 788 788 H -> N (IN P2 FORM).
 CC FT SEQUENCE 808 AA; 87567 MW; 0F4160DA78652445 CRC64;
 SQ
 Query Match 9.9%; Score 200.5; DB 1; Length 808;
 Best Local Similarity 20.9%; Pred. No. 6,4e-08;
 Matches 101; Conservative 61; Mismatches 173; Indels 149; Gaps 19;
 QY 1 IVANGVIV-AGSTOYSPFCFVSGHDSATGELMKNYFPRAGGEGDETWGNDY 55
 DB 349 ITANALIDNDSVKAQS--ATOAFDYITKRWVWDASPPNDLPDESHVYFHN 405
 QY 56 ARMTGAMGOITYPVNLVH--VGTAVGPASRGTGCTGTGNTFFAVAPDGEITW 115
 DB 406 NSWIVS-----SIDANLNLVIYPMGV-----GPDQGGSDRTKDSERFAGIALN 451
 QY 108 PDTEIVRHQTLPRDNDDECTFEMNTVDVQPTSTEMGLQSIENFNAATGERVLTG 167
 DB 452 ADTGKLAWFYQTVHDDMDLPSQPSLVDTOKDGLVPAIYA-----495
 QY 168 PCKTGTWQDAETG-EFLMARDT-----NYONMIESIDENIVVNE-----217
 DB 496 PRTQDITVLDRTVKGKPLRVIEPIKADIPGEYSKTPRSVEMQIQNQLKESDMWG 552
 QY 218 V-----EVD-----VCPFLGGRDPSALNDPSIYF 247

```

Db 553 IWGTTIDQMFCSITFHTLNEGFPPTSLKSLIPFGDLMGEWGLAVDPOVAFAN 612
QY 248 LNNVCYDMAVADQ-----EFTSMQVYNTSNTKLPKGM----- 283
Db 613 PISLPEFSQVLPRGPGNPLPEENAKGTGETGLQHNNGIPYAVNLHPFLDPLLPGLK 672
QY 284 -----GRDAIDISTGRTLSVERAANYS-----PV-----LSTG 314
Db 673 MPCRTPMGVYAGIDLTNNVWQHNRGTLRDSMYSLEPIPLPIKIGVPSLGPSTPA 732
QY 315 GGVLE-NGTDRYFRALSOETGETLMOTRLATVASGAQISYEVDGMQYVALIAGGYSYS 373
Db 733 GNGFLFASMDYITRAVNLTTGKVLMDRLPAGAQAATPIYATNGQYI-----VTYAG 786
QY 374 GLNS 377
Db 787 GHNS 790

RESULT 15
QY 001A_XANCU STANDARD; PRT; 790 AA.
AC 09XD78;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Probable quinolate dehydrogenase [Pyrroloquinoline-quinone]
   (EC 1.1.99.25).
GN OUMA.
OS Xanthomonas campestris (pv. Juglandis).
OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
OC Xanthomonas.
ON NCBI_TaxID=4291;
RX 11
RP SEQUENCE FROM N.A.
RC STRAIN-C5;
RX MEDLINE=20063481; PubMed=10594704;
RA Lee Y.-A., Lo Y.-C., Yu P.-P.;
RT "A gene involved in quinolate metabolism is specific to one DNA homology
   group of Xanthomonas campestris."
RL J. Appl. Microbiol. 87:649-658(1999).
CC -1- CATALYTIC ACTIVITY: Quinolate + pyrroloquinoline-quinone = 5-
   dehydroquinolate + reduced pyrroloquinoline-quinone.
CC -1- COFACTOR: PQO (BY SIMILARITY).
CC -1- PATHWAY: QUINIC ACID CATABOLIC PATHWAY; FIRST STEP. THIS PATHWAY
   ALLOWS GROWTH OF BACTERIA WITH QUINATE BY ITS CONVERSION TO
   PROTOCATECHUATE AND SUBSEQUENT METABOLISM BY THE BETA-KETOADIPATE
   PATHWAY.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL PQO DEHYDROGENASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF109471; AAC38453.1; -
DR InterPro; IPR001479; Bac_PQO.
DR InterPro; IPR002372; Bac_PQO_repeat.
DR Pfam; PF01011; Bacterial_PQO; 6.
DR PROSITE; PS00363; BACTERIAL_PQO_1; FALSE_NEG.
DR PROSITE; PS00364; BACTERIAL_PQO_2; FALSE_NEG.
KW Oxidoreductase; PQO; Quinolate metabolism; Transmembrane.
FT TRANSMEM 22 42 POTENTIAL.
FT TRANSMEM 48 68 POTENTIAL.
FT TRANSMEM 77 94 POTENTIAL.
FT TRANSMEM 106 126 POTENTIAL.
SQ SEQUENCE 790 AA; 82896 MW; 875F29B52A49FE6F CRC64;

```

```

Query Match      8.9%; Score 180.5; DB 1; Length 790;
Best Local Similarity 22.3%; Pred. No. 2.2e-06;
Matches 93; Conservative 43; Mismatches 150; Indels 131; Gaps 17;

QY 51 GNDYERARMTGAMGOITVDPTNLVHYSSTAVGASSETQRTPGTLYGTN----- 101
Db 386 GSSY-VRSTPNWAMPSYDAAMNTVF-----LPLGSPSTDLYGAEFTALDHRYG 433
QY 102 -RFAVRPDGTGEIVRHOITLPRDNWDQCTEFEMANTNDVPSFMEGLSINRNATGE 160
Db 434 ASVLLADLTGAEKVVQTFVHNDLMD-----FDLPQPSL-----IDPPNODGS 477
QY 161 RRVLTGVPCKTGTMMQFPAETGEFLW-ARDT-----NY---QNMIESIDENCIYT 206
Db 478 HTPAVVIGTKAGQIVLVDRATGKPLTEVREVPVAGSDIAHQVAPPTQPLSGMQITKH 537
QY 207 VNEDAILKELDVEYDVC-----PTFLGGRDPSAALNDPSG 242
Db 538 LTESDMGATAMDQMLCRIFAKOMREGLYTPGTDVLSLSPGSLGGMNMGSLSTDVHD 597
QY 243 IYF-----IPLNN-----VCYDMAVADQEFMSMDYNTSNTKLP--- 276
Db 598 VVFANDMRLGLMVOMITPADTRKAPADAGGAEAVNTGMGAVPLKGPYAVNKNRFLSALGIP 657
QY 277 ---PPGKDMIGRIDAIDISTGRTLSVERAAN-----YSPV-----LSTG 314
Db 658 CQAPP---YGTLSAIDLKTRSIAMQVPVGVQDTGSPFGIKMHLPIPIGMPTLGITLSTQ 713
QY 315 GGVLENGT-DRTFRALSQETGETLMOTRLATVASGAQISY--EVDGMQYVALIAGG 368
Db 714 GGLVFIAGTQDYLRLARDSATGKELMKGRPLVPGSGGPIYVSHKTKQYVVISAGG 770

```

Search completed: May 24, 2002, 10:15:52
 Job time: 770 sec

Fri May 24 11:27:11 2002

us-08-934-506a-5_copy_180_556.rsp

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2002, 10:17:32 ; Search time 93.93 Seconds

(without alignments)
694.337 Million cell updates/sec

Title: US-08-934-506a-5_COPY_180_556

Perfect score: 2032
Sequence: 1 IVANGVIVAGSTCOYSPFCG.....GMQYVIAAGGVSYSGSLINS 377

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_19:*
1: SP_archaea:*
2: SP_bacteria:*
3: SP_fungi:*
4: SP_human:*
5: SP_invertebrate:*
6: SP_mammal:*
7: SP_mmc:*
8: SP_organelle:*
9: SP_phage:*
10: SP_plant:*
11: SP_rodent:*
12: SP_virus:*
13: SP_vertebrate:*
14: SP_unclassified:*
15: SP_virus:*
16: SP_bacteriaph:*
17: SP_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1275.5	62.8	608	Q93RE9	Q93RE9 pseudogluc
2	394	19.4	698	Q9KH03	Q9KH03 alcaligenes
3	380.5	18.7	601	P71509	P71509 methylolact
4	378.5	18.6	629	Q9A048	Q9A048 methylolact
5	375	18.5	601	Q9EYWB	Q9EYWB rhizobium m
6	374	18.4	601	Q92WY9	Q92WY9 rhizobium m
7	360.5	17.7	695	Q9F9U2	Q9F9U2 pseudomonas
8	359	17.7	708	Q9L935	Q9L935 methylolovoru
9	355.5	17.5	708	Q9A644	Q9A644 comamonas t
10	350.5	17.2	633	Q24759	Q24759 hyphomicrob
11	338	16.6	742	Q53362	Q53362 acetobacter
12	335.5	16.5	573	Q95540	Q95540 methylolophil
13	323	15.9	695	Q934G0	Q934G0 pseudomonas
14	299.5	14.7	691	Q9AF95	Q9AF95 pseudomonas
15	281.5	13.9	623	Q9AGW3	Q9AGW3 pseudomonas
16	226.5	11.1	790	Q9X2S5	Q9X2S5 pantoea cit

17	208.5	10.3	182	2	Q92699	Q92699 hyphomicrob
18	204.5	10.1	470	2	Q90326	Q90326 acetobacter
19	203.5	10.0	179	2	Q92697	Q92697 hyphomicrob
20	200.5	9.9	180	2	Q92696	Q92696 hyphomicrob
21	200.5	9.9	180	2	Q92700	Q92700 hyphomicrob
22	200.5	9.9	181	2	Q92615	Q92615 hyphomicrob
23	199.5	9.8	181	2	Q92692	Q92692 hyphomicrob
24	197.5	9.7	181	2	Q92703	Q92703 hyphomicrob
25	196.5	9.7	181	2	Q92621	Q92621 hyphomicrob
26	195.5	9.6	182	2	Q92706	Q92706 hyphomicrob
27	193.5	9.5	180	2	Q92701	Q92701 hyphomicrob
28	191.5	9.4	182	2	Q92612	Q92612 hyphomicrob
29	191.5	9.4	184	2	Q92693	Q92693 hyphomicrob
30	190.5	9.4	172	2	Q93894	Q93894 hyphomicrob
31	186.5	9.2	172	2	Q93882	Q93882 methylolusinu
32	186.5	9.2	179	2	Q92707	Q92707 hyphomicrob
33	185.5	9.1	180	2	Q92694	Q92694 hyphomicrob
34	185.5	9.1	180	2	Q92704	Q92704 hyphomicrob
35	185	9.1	180	2	Q93884	Q93884 methylolusinu
36	183.5	9.0	171	2	Q92702	Q92702 hyphomicrob
37	182.5	9.0	184	2	Q92695	Q92695 hyphomicrob
38	181.5	8.9	185	2	Q92695	Q92695 uncultured
39	177	8.7	644	2	Q92651	Q92651 pseudomonas
40	176.5	8.7	172	2	Q93K58	Q93K58 methanotrop
41	175.5	8.6	172	2	Q93325	Q93325 methylolacell
42	174.5	8.6	179	2	Q92698	Q92698 hyphomicrob
43	174	8.6	803	16	Q91115	Q91115 pseudomonas
44	172.5	8.5	185	2	Q91810	Q91810 uncultured
45	169.5	8.3	185	2	Q918K9	Q918K9 uncultured

ALIGNMENTS

RESULT 1
Q93RE9 PRELIMINARY; PRT; 608 AA.
AC Q93RE9
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE ALCOHOL DEHYDROGENASE.
GN ADH.
OS Pseudoglucobacter saccharoketogenes.
OC Bacteria; Pseudoglucobacter.
OX NCBI_TaxID=133921;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-IFO 14464;
RA Shihata T., Saito Y.;
RT "Alcohol dehydrogenase."
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB046580; BAB62258.1; -
SQ SEQUENCE 608 AA; 65101 MW; 0ACBC97AE1BA570 CRC64;

Query Match 62.8%; Score 1275.5; DB 2; Length 608;

Best Local Similarity 60.4%; Pred. No. 2.6e-88;
Matches 236; Conservative 43; Mismatches 97; Indels 15; Gaps 3;

QY	1	IVANGVIVAGSTCOYSPFCGVSCHDSATGEBELMRNFIPRAGEDEETNGN-DYEARM 59
DB	207	IVVGVVIVAGSTCOYSPFCGVSCHDSATGEBELMRNFIPRAGEDEETNGAFAEYRM 266
QY	60	TGAGQITVDPVNLVHGSTAVGPASETORGTGGTLVGINTRFAVAPDGTGLVWRHQT 119
DB	267	TGAGQITVDELDLVYGGTAVGPASETORGTGGTLVGINTRFAVAPDGTGLVWRHQT 326
QY	120	LPRNMDOCTFFEMVATVNDVOSTEMEGLOSTIPNATGE-RVLTGVGCKTGTMQPD 178
DB	327	LPRNMDOCTFFEMVATVNDVOSTEMEGLOSTIPNATGE-RVLTGVGCKTGTMQPD 386
QY	179	AETGEPLMARDTNYQNMIESIDENGIVTVNEDATLKLDEYDVCPFTGLGRMPMAALN 238

```

Db 387 AKTGDFWMSKATVEONSISIDDTGLVTVEMDLKEPKETNYCPTFLGGRDMPAGYL 446
QY 239 PDGIFEIFPLNNCYDMAVDOEFMSDNYNTSNKLPKGMIGRIDALIDISTERTM 298
Db 447 PKSNLYIPLSNACVDMATTEATPADYNTDAILVLPKGMKMRDAIDLATGETKW 506
QY 299 SVERAANYSPLVSTGGVLFNGGIDRYFRALSOETGETLMQTRLATVASGOALSYEVND 358
Db 507 SYEFRAALYDPLVLTGGDLVFGVIGIDRFRALDAESGKSWRRLPGANSGYTSIDG 566
QY 359 MOVVALAGG-----VSYGSLN 376
Db 567 ROYVAAGVSGSLGGPTGPTTDPVDASGAN 597

RESULT 2
Q9KH03 PRELIMINARY; PRT; 698 AA.
AC 09KH03;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE TERAHYDROFURYL ALCOHOL DEHYDROGENASE.
OS Alcaligenes eutrophus (Ralstonia eutropha).
OC Bacteria; Proteobacteria; beta subdivision; Ralstonia group;
OC Ralstonia.
OX NCBI_Taxid=510;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21125557; Pubmed=11222593;
RA Zarrt G., Schrader T., Andreessen J.R.;
RT "Catalytic and Molecular Properties of the Quinolomoprotein
RT Tetrahydrofuryl Alcohol Dehydrogenase from Ralstonia eutropha
RT Strain Bo."
RL J. Bacteriol. 183:1954-1960(2001).
DR EMBL: AF273737; AAF6335.1; -.
DR HSSP; Q92477; IFIG.
DR InterPro; IPR002372; Bac_PQQ_repeat.
DR InterPro; IPR003045; Cylc_heme_bind.
DR Pfam; PF01011; Bacterial_PQQ; 6.
DR Pfam; PF00034; Cytochrome_C; 1.
DR PROSITE; PS00190; CYTOCHROME_C; UNKNOWN.1.
SQ SEQUENCE 698 AA; 75471 MW; 9C87F180BD12FB2 CRC64;

Query Match 19.4%; Score 394; DB 2; Length 698;
Best Local Similarity 28.7%; Pred. No. 1.6e-21;
Matches 117; Conservative 61; Mismatches 161; Indels 68; Gaps 17;

QY 2 VANG-VTVAGSTCOYSPFGCVSGHDSATGELMKNRYIP-----RAGEEGDETWG 51
Db 192 VYNGKVIIGNGAGAEYGRS-YITAYDAETGKQOMRWIYVPGPARFENEMAKAAATW- 249
QY 52 NNYEAR-WMTG-----AMGQITDPTVNLVHGSTAVGPASSETORCTPGG-TLYGTNTPFA 105
Db 250 -DPSGRYVINGGGGTVMWMAFDPELNLMIYIGTGNAGPWSRKLRSRPGKMDMLYAASTV-VA 307
QY 106 VAPDGEIYWRHQTLPBPNWMOECFEEMVYN--VDVQPSMEMELOSINNAATGERKY 163
Db 308 LMPDGEIYVWHTQETPGDMDWTTSTODITLLADLKIDGDP-----RKV 349
QY 164 LTVGPKCTGTMMQPDATGEFLMARDTNVQNMIIESIDENG--IYVNEADALIKELDEYD 221
Db 350 IHAAR-KNGFEVIDRTNGKFIKSAKNFVDVWMAKSGDKNGRPVETPROADTSGKRA-----D 404
QY 222 VCPPLGRDMPAALNPDGSIYFTPLNNVCYDMAVDOEFMSDNYNTSN----- 272
Db 405 VYVGPFGAHHMSMFHFKLGIAPQHPPL-TLADNKEW-----VHNQKDSPEAHNGVG 459
QY 273 -----VTKLPKGMIGRIDALIDISTGRTLSVERAANYSPLVSTGGVLFNGGIDRY 326

```

```

Db 460 WNLGLVNAEPPRSKPMGRLLAMPDLAQKAVHHRDHAGPNNGTLATAGNLVFGGTADGR 519
QY 327 FRALSOETGETLMQTRLATVASGOALSYEVNDQVYALV-GGVSYG 372
Db 520 LVVAHAATGEKLMQAPGSGVVAAPVYLLDGRQYVAVAGMGVYG 566

RESULT 3
P71509 PRELIMINARY; PRT; 601 AA.
AC P71509;
DT 01-FEB-1997 (Tremblrel. 02, Created)
DT 01-FEB-1997 (Tremblrel. 02, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE METHANOL DEHYDROGENASE LARGE SUBUNIT HOMOLOG.
GN MXPAP.
OS Methylobacterium extorquens.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Methylobacterium group; Methylobacterium.
OX NCBI_Taxid=408;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AM1;
RX MEDLINE=97312011; Pubmed=9166622;
RA Christoserdova L., Lidstrom M.E.;
RT "Molecular and mutational analysis of a DNA region separating two
RT methylothrophy gene clusters in Methylobacterium extorquens AM1."
RL Microbiology 143:1729-1736(1997).
DR EMBL; U72662; AAB58890.1; -.
DR HSSP; P38539; AAAH.
DR InterPro; IPR002372; Bac_PQQ_repeat.
DR Pfam; PF01011; Bacterial_PQQ; 7.
SQ SEQUENCE 601 AA; 64952 MW; 68E45C7059CB8239 CRC64;

Query Match 18.7%; Score 380.5; DB 2; Length 601;
Best Local Similarity 27.1%; Pred. No. 1.4e-20;
Matches 108; Conservative 71; Mismatches 156; Indels 63; Gaps 15;

QY 18 FG-CFVSGHDSATGELMKNRYIP-----PRAGEEGDETWGNDYEAR 57
Db 193 FGVGCHTAIDLKSGKVMKGYISIGPPDQILVDEPKTSLGKPLAKDSILKTEGD---Q 249
QY 58 WMTG---AMGQITDPTVNLVHGSTAVGPASSETORCTPGGTLTYGTNTPRAVPTGETY 114
Db 250 WKTGGCGTWGWFSDYDPLDLMTYVSG--NPSTNMPKQRPQDNKW-SWTIARBPDIQMAK 306
QY 115 WRHQTLPDNDWDOCTEEMAVTNDVOPSTEMEGLSINNAATGERRVLTGVPCKTGTW 174
Db 307 WVYQMTPHDEDFDGINEMILT-----OKFDG-----KDRPLTHFD-RNGFG 349
QY 175 WQFAEFGELMARDTN-YQNMIESID-ENGIYVNEADALIKELDEYD-----VCP 224
Db 350 YTLDRATGEVLVAEKFPVYNMATKYDLDGKSTYGRPLVYSYSTEDONKEDVNSGICP 409
QY 225 TFLGGRDMPAALNPDGSIYFTPLNNVCYDMAVDOEFMSDNYNTSNYTKLP-PGK-DM 282
Db 410 AALGTAKQQAAPASPKGLFVYPTLNHCMDYEFERLYTPGQPYVATLSMYPAPSSHG 469
QY 283 IGRIDAIDISTGRTLSVERAANYSPLVSTGGVLFNGGIDRYFRALSOETGETLMQTR 342
Db 470 MGNFIADNDLQGIKISNPNQFSAMGALATSGDVLFTGLEGFLKAVDSKTKGKELYRK 529
QY 343 LATVASGOALSYEVNDQVYALV-----GVSYGSL 375
Db 530 TPGSIIGNVMTYEBHKQKQHAVLVSGVAGIGLAAGL 567

```

```

RESULT 4
Q9AQ48 PRELIMINARY; PRT; 629 AA.
AC Q9AQ48;
DT 01-JUN-2001 (Tremblrel. 17, Created)

```


RT fixing endosymbiont *Sinorhizobium meliloti*;
 RL Proc. Natl. Acad. Sci. U.S.A. 98:9889-9894(2001).
 DR EMBL: AL603642; CAC48573.1; -
 KW Oxidoreductase; Plasmid; Hypothetical protein; Complete proteome.
 SO SEQUENCE 601 AA; 65759 MW; D73424EFBD15ADBE CRC64;

Query Match 18.4%; Score 374; DB 16; Length 601;
 Best Local Similarity 26.5%; Pred. No. 4.3e-20;
 Matches 113; Conservative 66; Mismatches 161; Indels 86; Gaps 15;

QY 2 VANGVIVAGSTCOYSPFGCVSGHDSATGEELMRYFT-----PRAG 43
 DB 181 VKDKILVIGSGEFGVGRG-HYTAYSMAADGKVLKRGYSKGPDSDLIDPEKTHLKGKPYGK 239
 QY 44 EGGDETMGNDYEAARMGTG---AMGQITDPTNLVHGSTAVGASSTOGRGTGTLXGT 100
 DB 240 DSGLTITWED---QWKIGGGTTMGWYSTDPEENLVYGTGPNSTWNPQR--PGDNRW-S 293
 QY 101 NTRAVRPDTEIVRHQTLPRDMDQCTFEEMVTVNDVOPSTEMEGLOSINPNATGE 160
 DB 294 MTFIARDVDYGMKAVLYQMTPEHWDYDGVNEMILTEQDIDGK-----D 337
 QY 161 RRVLTGVCKGTMMQFAETGEFL-----WADTNYQNKIE-----SI 199
 DB 338 RKLTHFD-RNGFGYTMRYVTEGELLVAEKYDPTVMATEVMDKSDYGRPYVAQYST 396
 QY 200 DENGIVTVNEDAILKELDEYDVCPTFLGGRDWPSSALNPDGSIYFIPLNNVCYDMAVD 259
 DB 397 EQNG-----EDTNTT-----GCPALGTDQDQPAATSPKTELFYPTNHVCMDYEPFR 445
 QY 260 QEFTSMDEVNTSNVTKLPKGDW---IGRIDAIDISTGRILMSVERAANYSPVLSTGG 316
 DB 446 VSYTAGOPYVATISMYP-KDSHGKMGNFAMDNKEGKIKWSPPEFVSWSGALATAGD 504
 QY 317 VLFNGTDRYFRALSGEETELMOTRLATVASGQAISEYDGMQYVALIAG-----GVSY 371
 DB 505 VVFYGTLEGIYKAVDATGKELYRFPKTPSGVIGVMTYARAGKYVALVSGVGMAIGL 564
 QY 372 GSGINS 377
 DB 565 AAGLTN 570
 RESULT 7
 Q9F902 PRELIMINARY; PRT; 695 AA.
 AC Q9F902;
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE ALCOHOL DEHYDROGENASE.
 GN ACH.
 OS *Pseudomonas stutzeri* (*Pseudomonas perfectomarina*).
 OC Bacteria; Proteobacteria; gamma subdivision; *Pseudomonadaceae*.
 OC *Pseudomonas*.
 OX NCBI_Taxid=316;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BC-2;
 RA Chang C.-H., Herrick J.B., Okinaka R.T., Brainard J.B.,
 RA Terwilliger T.C.;
 RT "Identification and characterization of genes activated by 2-
 RT chloroethanol in *Pseudomonas stutzeri* BC-2.";
 RT Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF176640; AAG09249.1; -
 DR HSSP: Q9247; 1FLG.
 DR InterPro: IPR002372; Bac_PQO_repeat.
 DR InterPro: IPR003085; Cyt_PQO.
 DR InterPro: IPR003088; Cyt_CT.
 DR Pfam: PF01011; Bacterial_PQO; 6.
 DR Pfam: PF00034; Cytochrome_C; 1.
 DR PROSITE: PS00190; CYTOCHROME_C; UNKNOWN_1.

SO SEQUENCE 695 AA; 75842 MW; D29A698A642E1A83 CRC64;

Query Match 17.7%; Score 360.5; DB 2; Length 695;
 Best Local Similarity 27.9%; Pred. No. 5.6e-19;
 Matches 109; Conservative 57; Mismatches 173; Indels 51; Gaps 14;

QY 1 IVANGVIVAGSTCOYSPFGCVSGHDSATGEELMRYFIPRAGE-----BGDET 49
 DB 185 VVKGKVLITNGCAEYVGRG-FFSAYDAETGKMMARFYTP--GDPAPQYEPHPLAAMKT 241
 QY 50 W-GNDYEAARMGTG---AMGQITDPTNLVHGSTAVGASSTOGRGTGTLXGTNTR 104
 DB 242 WKGDQY---WKLGGGVTWPMKAYDELDLITGTGSGWNEIRSPGGDNLVLSIL 298
 QY 105 AVRPDTEIWRHQTLPRDMDQCTFEEMVTVNDVOPSTEMEGLOSINPNATGERVY 164
 DB 299 ALRPDSGKLMHYQTMPEHWDYDGVNEMILTEQDIDGK-----KRRVL 342
 QY 165 TGVPCKGTMMQFAETGEFLARDINQNMISID--ENGIVTVNEDAILKELDEYDV 222
 DB 343 MQAP-KNGFFYVLDRAITGELLSEKFGKYLMAEKVDLATGRPEYFGSRERKQVVM-- 399
 QY 223 CPTFLGGRDWPSSALNPDGSIYFIPLNNVCYDMAVDQETSMDEVNT---SNVTKLP 278
 DB 400 -PSSFGANNHMSFNPQGTGMATIRYQELIPGYRNAGATFKKIDGNTGTSFSDHEIP- 457
 QY 279 GKMT-GRIDAIDISTGRILMSVERAANYSPVLSTGGVLFNGTDRYFRALSGEET 337
 DB 458 -RDVSGALILAMPVQRQREMRVPHSFVWNGGTLSTAGNLVPGVADGGLHAYSADKQGR 516
 QY 338 LMOTRLATVASGQAISEYDGMQYVALIAG 367
 DB 517 LMSFAQTGIVAAPISFSLDGEQYAVVMAG 546

RESULT 8
 Q9L935 PRELIMINARY; PRT; 599 AA.
 AC Q9L935;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE METHANOL DEHYDROGENASE LARGE SUBUNIT.
 GN MXAF.
 OS *Methylovorus* sp. (strain SSI / DSM 11726).
 OC Bacteria; Proteobacteria; delta subdivision; *Methylophilus* group;
 OC *Methylovorus*.
 OX NCBI_Taxid=81683;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-SSI;
 RA Kim Y.M.;
 RT "Cloning and nucleotide sequence of mxaf gene of *Methylovorus* sp.
 RT strain SSI DSM11726.";
 RT Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF184915; AAD56237.2; -
 DR HSSP: P38539; 4AAH.
 DR InterPro: IPR001479; Bac_PQO.
 DR InterPro: IPR002372; Bac_PQO_repeat.
 DR Pfam: PF01011; Bacterial_PQO; 7.
 DR PROSITE: PS00364; BACTERIAL_PQO_2; 1.
 SO SEQUENCE 599 AA; 65133 MW; DBF6F4B5D871BC91 CRC64;

Query Match 17.7%; Score 359; DB 2; Length 599;
 Best Local Similarity 25.9%; Pred. No. 5.9e-19;
 Matches 106; Conservative 70; Mismatches 161; Indels 72; Gaps 14;

QY 10 GSTQGYSPF-----GC-----FVSGHDSATGEELMRYFT----- 39
 DB 177 GSTTQAFPAVKVYLVGCGAELGVRGYVATYADQKGTGLVWRSFATGPDEINLAKDFN 236

QY 40 ---PRAGEE---GDETWGNDYEAARMTGA---WQITVDPVTLNLYGSTAVGASSETORG 91
 DB 233 KNPFTGQMGJSTKTWESE---GKIGGGTMMGWYAYVPRKLNLFYSGGNAPAWMETMR- 292
 QY 92 TPEGGLYGTNFRFAVRPTGTETVWHQTLPRDNWDECTEFEMATNVVQSTEMEGIOS 151
 DB 293 -GDNKRW-TMTIWARDVDTGAARKWGYOQTPHDEMFPAGVNDMLTDQVNGKTO----- 344
 QY 152 INPMNATGRRVLTGVPCQCTGTMQFDETFEFLWARDTN-YQNMIESIDENGITVNEED 210
 DB 345 -----PLTHVND-RNGIMTILNQTGSIVQAADVAVNFKVVDLKTGLPVPDP 393
 QY 211 AILKEIDVE-YVVCPTFLGGRDPSAALNPDSCGIFIPLNVCYIMAAVDEFTSMAYN 269
 DB 394 ESTRDHDKGTNCPASAMFHNQGLDAYDPDSRTFFYGLNHCMDMEPFMLPYRAGQFEV 453
 QY 270 TSNVTKLP----PKDIMGRIADIDISTGRTLSYERAAANSPLYSTGGVLENGTDR 325
 DB 454 GATLAMPGPNGPTKEMQVLAQVGTGEVYKTKRKESVNGGLTATGGLVFNITLDG 513
 QY 326 YFRALSGETGLMOTRLATVASGQAISEVDGMQVYALAGGVSTGSG 374
 DB 514 NIKALDKTKNKELMKFKMPSGALGAPMSYAYKKGXYIA-----TNGVG 557

RESULT 9
 Q46444 ID 046444 PRELIMINARY; PRT; 708 AA.
 AC 01-NOV-1998 (TEMBREL. 08, Created)
 DT 01-NOV-1998 (TEMBREL. 08, last sequence update)
 DT 01-DEC-2001 (TEMBREL. 19, last annotation update)
 DE QUINOHAEPROTEIN ETHANOL DEHYDROGENASE TYPE I PRECURSOR
 DE (EC 1.1.99.-) (OH-EDH).
 GN OHEDH.
 OS Comamonas testosteroni (Pseudomonas testosteroni).
 OC Bacteria; Proteobacteria; beta subdivision; Comamonadaceae; Comamonas.
 OX NCBI_TaxID=285;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 15667;
 RX MEDLINE=96184549; PubMed=8654419;
 RA Stoorvogel J., Kraayveld D.E., van Sluis C.A., Jongejan J.A.,
 De Vries S., Duine J.A.;
 RT "Characterization of the gene encoding quinochaemoprotein ethanol
 dehydrogenase of Comamonas testosteroni";
 RL Eur. J. Biochem. 235:690-698(1996).
 RN [2]
 RP SEQUENCE OF 32-54 AND 477-490.
 RC STRAIN=ATCC 15667;
 RX MEDLINE=95324580; PubMed=7601151;
 RA De Jong G.A.H., Geerlof A., Stoorvogel J., Jongejan J.A., De Vries S.,
 Duine J.A.;
 RT "Quinochaemoprotein ethanol dehydrogenase from Comamonas testosteroni.
 Purification, characterization, and reconstitution of the apoenzyme
 with pyrroloquinoline quinone analogues";
 RL Eur. J. Biochem. 230:899-905(1996).
 RN [3]
 RP CHARACTERIZATION.
 RX MEDLINE=86242113; PubMed=3521592;
 RA Groen B.W., van Kleef M.A., Duine J.A.;
 RT "Quinochaemoprotein alcohol dehydrogenase apoenzyme from Pseudomonas
 testosteroni";
 RL Biochem. J. 234:611-615(1986).
 RN [4]
 RP CRYSTALLIZATION.
 RX MEDLINE=21536088; PubMed=11679760;
 RA Oudrie A., Huijzinga E.G., Rozeboom H.J., Kalk K.H., de Jong G.A.H.,
 Duine J.A., Dijkstra B.W.;
 RT "Crystallization of quinochaemoprotein alcohol dehydrogenase from
 Comamonas testosteroni: crystals with unique optical properties";
 RL Acta Crystallogr. D 57:1732-1734(2001).
 CC -1- FUNCTION: CATALYZES THE DYE-LINKED OXIDATION OF PRIMARY ALCOHOLS

CC TO THE CORRESPONDING ALDEHYDES AND THE (SUBSEQUENT) OXIDATION OF
 CC THE ALDEHYDES TO CARBOXYLIC ACIDS.
 CC -1- COFACTOR: POQ, HEME, AND CALCIUM.
 CC -1- SUBUNIT: MONOMER.
 CC -1- SUBCELLULAR LOCATION: PERIPLASMIC (POTENTIAL).
 CC -1- INDUCTION: BY ETHANOL AND BUTANOL.
 CC -1- SIMILARITY: TO ALCOHOL DEHYDROGENASE (ACCEPTOR) PROTEINS.
 DR EMBL: X81880; CA57464.1; -
 DR HSSP: Q924J7; 1F1G.
 DR Interpro: IPR002372; Bac_POQ_repeat.
 DR Interpro: IPR003086; Cyt_C1.
 DR Interpro: IPR002329; Cyt_C1C.
 DR Pfam: PF01011; Bacterial_POQ; 6.
 DR Pfam: PF00034; cytochrome_c; 1.
 DR PRINTS: PR00605; CYTCROME.C.
 KW Signal; POQ; Heme; Calcium; Oxidoreductase; Periplasmic.
 FT SIGNAL 1 31
 FT CHAIN 32 708
 FT BINDING 635 635 QUINOHAEPROTEIN ETHANOL DEHYDROGENASE
 FT BINDING 638 638 TYPE 1.
 FT METAL 639 639 HEME (COVALENT) (BY SIMILARITY).
 FT METAL 639 639 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
 SQ SEQUENCE 708 AA; 76822 MW; 99AB54BDD6ACAB3 CRC64;

Query Match 17.5%; Score 355.5; DB 2; Length 708;
 Best Local Similarity 26.7%; Pred. No. 1.4e-18;
 Matches 105; Conservative 59; Mismatches 180; Indels 49; Gaps 12;

QY 6 VYVASTQYSPFGCGVGHDSATGEELMRNRYFP-----RAGEGDETWGNDYEA 56
 DB 208 VILGRGALEYGRG-VITATDAETGERKRMFSYVGDSKPFEDSMKRAARTW--DPSG 264
 QY 57 RMTGA-----WQITVDPVTLNLYGSTAVGASSETORGTPGGTLTYNTRFAVRPTG 111
 DB 265 KMWEGGGGTMDSMTPFDALNTMYVGTGNGSPWGHKRSFGNDLTLASTVALADPTG 324
 QY 112 EIVWRHQTLPKRNWQCECFEMATNVVQSTEMEGIOSINPMNATGRRVLTGVPCKT 171
 DB 325 KYRMHQTETPGDNWDYTSQPMILADIKI-----AGRPVYILNAP-KN 367
 QY 172 GTMMQFDETFEFLWARDTNQNMIESIDENG-IYVNEDAILKELDYEDVCPFLGGR 230
 DB 368 GFPEFVLDNRNGKFKISKNNVPVNMASGYDKHGKPIGI--AAARGSKRQDAPVPGYGAH 424
 QY 231 DWPSAALNPDSCGIFIPLNVCYDMAAYDO-EFT-----SMDVYNSVTKL-PPGK 280
 DB 425 NMHPMSFNPQTGLVLYPAQNVPNVLDKDKWENQAGPQKPOSGTCGMNTAKFFNAPRKS 484
 QY 281 DWIGRIDAIDISTGRTLSYERAAANSPLYSTGGVLENGTDRYFRALSGETGLMIO 340
 DB 485 KPEGRLLAMPVQAQKAMSEVHSPNNGGTLTLAGNVFQGTADRLVAYHNATGEKIME 544
 QY 341 TRLATVASGQAISEVDGMQVYALAGGVSTG 372
 DB 545 APTGTGVAAPSTYVMDGRQYVSAVAGWGCVYG 577
 RESULT 10
 Q24759 ID 024759 PRELIMINARY; PRT; 633 AA.
 AC 024759;
 DT 01-JAN-1998 (TEMBREL. 05, Created)
 DT 01-JAN-1998 (TEMBREL. 05, last sequence update)
 DT 01-DEC-2001 (TEMBREL. 19, last annotation update)
 DE METHANOL DEHYDROGENASE ALPHA SUBUNIT PRECURSOR (EC 1.1.99.8).
 DE KXAF.
 GN Hyphomicrobium methylovorum.
 OS Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Hyphomicrobium group; Hyphomicrobium.
 OX NCBI_TaxID=84;
 RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN-GM2;
MEDLINE-97457202; PubMed-9311140;
RA Tanaka Y., Yoshida T., Matanabe K., Izumi Y., Mitsuura T.;
RT "Cloning and analysis of methanol oxidation genes in the methylotroph
RT *Hyphomicrobium methylovorum* GM2.";
RL FEMS Microbiol. Lett. 154:397-401(1997).
DR EMBL: AB004097; BAA23272.1; -.
DR HSSP: P38539; 4AAH.
DR InterPro: IPR001472; Bac_PQQ.
DR Pfam: PF01011; Bacterial_PQQ.
DR PROSITE: PS00363; BACTERIAL_PQQ_1; 1.
DR PROSITE: PS00364; BACTERIAL_PQQ_2; 1.
KW Signal; Oxidoreductase.
FT SIGNAL 1 34
FT CHAIN 35 633 POTENTIAL.
FT SEQUENCE 633 AA; 69853 MW; B47A23A2779E2C1B CRC64.

Query Match 17.2%; Score 350.5; DB 2; Length 633;
Best Local Similarity 26.2%; Pred. No. 2.8e-18;
Matches 111; Conservative 70; Mismatches 169; Indels 73; Gaps 16;

QY 2 VANGVIVAGSTQYSPFGC-----VSGHDSA-----TGEELMKNYFI----- 39
DB 181 VENGDIKYGQILQAPYVYHDAIYSSGAEIYAGHVAIVNKTGEQAMRYVATGPDEE 240
QY 40 -----PRAGEE--GDETMGNDYEAAMTGA--WGQITYPVTNLVHYGSTAVG 83
DB 241 IGLADDFNSANPHYGKGLGATMEGD---AMKIGGGTNGMAYADPQANLIYSSGSPA 297
QY 84 PASERORTPGGLTGTWTRFAVRPTGEIYMRHQTLPDNDWDECTEEMATVNDVOPS 143
DB 298 PWNEMNR--PGDNKW-TTITARADYGEKMGFGYOKTPHDEMDAGVNVIMLSE-----Q 349
QY 144 TEMEGLOSINPNAATGERRVLTGVPCKTGMQFPAETGEELMA---RDINYOAMIESID 200
DB 350 TDKE-----GKKRKLTLHPDRNGIYITLDRNGDISDKDDT--VAVFKNVD 396
QY 201 ENGIYTVNEDALIKELIYE-YDVCPTFLGGKMPSSALNDPSGIIYFIPLNNVCYDMAVD 259
DB 397 LKSGLPVPRDPEFTRMDBKHGTEICPSAMGYHNOGHDSDYPTKOLFEMGINHICMDMPEM 456
QY 260 QEFTSMDEVYNTSNTKLP-PGKDM-----IGRIDAIDISGRILMSVERAANYSPVLS 313
DB 457 LPRAGQEFVCAITLMTPGPKGRDQNLGLGQIKAYNAITNTYKWEHMERFSVWGTLAT 516
QY 314 GCGVLENGGIDRFRALSOETGETLMQTRLATVAGSQAISYEVDGMQYVAIAG-GVSYG 372
DB 517 AGNLVFEYGLDGLKARNSDTGLMKRLPSGVITGYPMTEYHKGYOYIAVMSGVGMPG 576
QY 373 SGL 375
DB 577 VGL 579

RESULT 11
ID 053362 PRELIMINARY; PRT; 742 AA.
AC 053362; Q44159;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-OCT-2001 (TREMBLrel. 18, Last annotation update)
DE DEHYDROGENASE SUBUNIT OF ALCOHOL DEHYDROGENASE.
OS Acetobacter pasteurianus (Acetobacter turbidans).
OC Bacteria; Proteobacteria; alpha subdivision; Acetobacteraceae;
OC Acetobacter.
NCBI_TaxID-438;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NC11380;
RX MEDLINE-94042848; PubMed-8226628;
RA Takemura H., Kondo K., Horinouchi S., Beppu T.;

RT "Induction by ethanol of alcohol dehydrogenase activity in Acetobacter
RT pasteurianus.";
RL J. Bacteriol. 175:6857-6866(1993).
DR EMBL: D13893; BAA40252.1; -.
DR HSSP: Q92407; 1FG.
DR InterPro: IPR001479; Bac_PQQ.
DR Pfam: PF01011; Bacterial_PQQ.
DR PROSITE: PS00363; BACTERIAL_PQQ_1; 1.
DR PROSITE: PS00364; BACTERIAL_PQQ_2; 1.
KW Signal; Oxidoreductase.
FT SIGNAL 1 34
FT CHAIN 35 633 POTENTIAL.
FT SEQUENCE 633 AA; 69853 MW; B47A23A2779E2C1B CRC64.

Query Match 16.6%; Score 338; DB 2; Length 742;
Best Local Similarity 26.7%; Pred. No. 3.1e-17;
Matches 114; Conservative 62; Mismatches 165; Indels 86; Gaps 16;

QY 2 VANGVIVAGSTQYSPFGC--FVSGHDSATGEELMKNYFI--PRAGEGD----- 47
DB 203 VAKGLVILGN--GSGEARGFVSAPFAETGKLMKRYTPNNKNEPDHADVADNVLMSKA 260
QY 48 -ETWGNDEYARMT-----GAMQITYPVTNLVHYGSTAVGPASETORTPGTLY--- 98
DB 261 YKTWGP--KGAWRQGGGGTVDLSLYDPSDLIY--LAVG-----NGSPNNYRSE 309
QY 99 -GTN-----TRFAVRPTGEIYMRHQTLPDNDWDECTEEMATVNDVOPSTMEGLOSI 152
DB 310 GIGSNPLFSLIVALKPETGEYVNHFOATPDMDYNSVOOIMTLDMV----- 357
QY 153 NPNAAAGE--RRVLTGVPCKTGMQFPAETGEELMA--RDINYOAMIESIDE--NGIYTVNED 210
DB 358 -----NGEMRHVYIWHAP--KNGFEYVLDATGEELAKNNVYQWANGDPLTGRPIYND 411
QY 211 AIIEKELDEVDVCEPTFLGGKMPSSALNDPSGIIYFIPLNNVCY-----DMA 257
DB 412 GLTYLTGKFTWYGIPLGAGNFMGAYSPKTHLVYPAHQIFGKKNQYGFKPPDAWN 471
QY 258 VQDEFTSMDEVYNTSNTKLP-PGKDM-----IGRIDAIDISGRILMSVERAANYSPVLS 317
DB 472 VGLDMKNKLPDPE-ARAYIKDLHGLLAMPVKMETWIKDKPGWNGVLTGGDL 530
QY 318 LKSGIDRFRALSOETGETLMQTRLATVAGSQAISYEVDGMQYVAIAG-GVSYG 367
DB 531 LFOGLANGEFHAYDATINGSDLYKFAQSGIITASPTYSVNGQYVAEVGNGIYPIG 590
QY 368 GVSYS 374
DB 591 GVGRTSG 597

RESULT 12
ID 059540 PRELIMINARY; PRT; 573 AA.
AC 059540;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE METHANOL DEHYDROGENASE HEAVY SUBUNIT (EC 1.1.99.8).
OS Methylophilus methylotrophus (Bacterium W3A1).
OC Bacteria; Proteobacteria; beta subdivision; Methylophilus group;
OC Methylophilus.
NCBI_TaxID-17;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=W3A1;
RA Xia Z.-X., Dai W.-W., Zhang Y.-F., He Y.-N., White S.A., Boyd G.D.,
RT Mathews F.S.;
RT "Determination of the Gene Sequence and the Three-dimensional
RT Structure at 2.4 Å resolution of Methanol Dehydrogenase from
RT Methylophilus methylotrophus.";
RL Submitted (NOV-1995) to the EMBL/Genbank/DBJ databases.
DR EMBL: U41040; AAA83765.1; -.
DR HSSP: P38539; 4AAH.

DR InterPro: IPR001479; Bac_P00.
 DR InterPro: IPR002372; Bac_P00_repeat.
 DR Pfam: PF01011; Bacterial_P00_7.
 DR PROSITE: PS00364; BACTERIAL_P00_2; 1.
 DR Oxidoreductase.
 KW SEQUENCE 573 AA; 62635 MW; A06CB3091BB8F0C CRC64;

Query Match 16.5%; Score 335.5; DB 2; Length 573;
 Best Local Similarity 24.5%; Pred. No. 3.3e-17;
 Matches 108; Conservative 65; Mismatches 166; Indels 101; Gaps 16;

3 ANGVVA-----GSGCYSPF-----GC-----FVSGHD 26
 DB 125 ANGHLLADANTGKIMWEVCDPRVSTLTQAPFVAKDTYLMGSGAELGVRVAVAFD 184
 QY 27 SATGEELR-----NYFIPRAGE--EGDETGNDEYEAAMTGA--WGQ 65
 DB 165 LKTGELKRAAFATGSDSVRLAKDFNSANPHYGGEGLGCTKWECD--AMKIGGCTMNGW 241
 QY 66 ITYDPTNLVHYGSTAVGPASGTORGTCG-----TLKGTNTRAVRPDGTETVWRHOTL 120
 DB 242 YAYDKLNLFFYSGNPNAPWNETMR--PGDNKMTWTWGRDL-----DTGAKMGTOKT 293
 QY 121 PRDNMDQECTFEFMAVTVNDVQSTMEGLQSIINPAATGERRVLTGVPCKTGTMMQFPAE 180
 DB 294 PHDEWDFAGVNOAVLTD-----QPVN-----GKMTPLSLHIDRNGILYTLNRE 336
 QY 181 TGEFLWARDTN-YONMTESIDENGLIVYNEADILKELDVE-YDVCPITLGGHWPMSALN 238
 DB 337 NGNLIYAEKVDPAVNVFKVDLTGTTPVDPFATRMHKGNTNIPSAAGFINGQVDSYD 396
 QY 239 PSQGIYFIFLNVCYDMAVNDQETSMOYNTSNVTKLP-----PGKDMIGRIDAIDISTG 294
 DB 397 PESRTLYAGLNHICMWEPPMLPYRAGOFFVCATLTAMTGPMPKKEKGQIRADILDTG 456
 QY 295 RLMSVERAANYSPVLTSGGVLFGNGTDRYFRALSOETGETLMOTRLATVAGQAISY 354
 DB 457 KAKMTWEKEFAAMGTLTKGGLWYATLDYGLAKLDNKGRELWNFKMPSGGISGSPMTY 516
 QY 355 EVDGMQYVAIAGGVSYSYG 374
 DB 517 SFKGYI-----GSMYGVG 531

RESULT 13
 Q934G0 PRELIMINARY; PRT: 695 AA.
 ID Q934G0;
 AC Q934G0;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE LUPANINE HYDROXYLASE PRECURSOR.
 GN LDI.
 OS Pseudomonas sp. DH2001.
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 OC Pseudomonas.
 OX NCBL_TaxID=163360;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Hopper D.J., Kaderbhai M.A., Little A.R., Marriott S.A., Young M.,
 RA Rogozinski J.,
 RA "Cloning, sequencing and analysis of the gene for lupanine
 RT hydroxylase, a quinuoytochrome c from a pseudomonas sp.";
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AJ118095; CAC67410.1; -
 KM Signal.
 FT SIGNAL 1 26
 FT CHAIN 27 695 LUPANINE HYDROXYLASE.
 FT SEQUENCE 695 AA; 74956 MW; 9096C6387E457FE0 CRC64;

Query Match 15.9%; Score 323; DB 2; Length 695;

Best Local Similarity 27.2%; Pred. No. 3.8e-16;
 Matches 101; Conservative 56; Mismatches 157; Indels 58; Gaps 13;

QY 24 GH-----DSATGEELMNYFIPRAGE-----EGDETGNDEYEAAMTGA--WGQ 63
 DB 202 GHDAFDARTGKILWFRPYTWP--GDPSEPPENDLAKASTKWTGTDY--WKTGKGVSPW 256
 QY 64 GQITDPTNLVHYGSTAVGPASGTORGTCGTLTGNTNTRAVRPDGTETVWRHOTL 123
 DB 257 DAITYEASPTLFGVDGSPSPAPQADACDELSHIAVASTGA YKMHFFQVND 316
 QY 124 NMDQECTFEFMAVTVNDVQSTMEGLQSIINPAATGERRVLTGVPCKTGTMMQFPAE 183
 DB 317 GSNMSATMHLMLDLDP-----EGV-----SKRYVMTAP-KNGTFYLDASTK 359
 QY 184 FLNARDTN YONMTESID-ENGLIVYNEADILKELDVEYDV-CPFLGGHWPMSALN 241
 DB 360 FISADHIVPNNWTKGDKPRKGRIPSEANAYWRRPGEMTIFLPGDVGHGHWEMAAVNP 419
 QY 242 GIYFIPLNNVCYDMA-VDOETSMOY-----NTSNVTKLPKPKMIGRIDAIDISTG 295
 DB 420 RYVIFPSTLVVYVAVASKDTGELDLDY YGMRDPAITKTG-----GDLVAMDPLOK 471
 QY 296 TLMSVERAANYSPVLTSGGVLFGNGTDRYFRALSOETGETLMOTRLATVAGQAISY 355
 DB 472 EKWAKRSLPVGVLATAGGLVFGTGDFEAFDANTEKLSRHYGSIILAPTYE 531
 QY 356 VDGMOYVAIAGG 367
 DB 532 VDGQYILVAVSG 543

RESULT 14
 Q9AF95 PRELIMINARY; PRT: 691 AA.
 ID Q9AF95;
 AC Q9AF95;
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE 1-BUTANOL DEHYDROGENASE BDH.
 DE Pseudomonas butanovora.
 OS Pseudomonas butanovora.
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 OC Pseudomonas.
 OX NCBL_TaxID=86174;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Vargdal A.S., Arp D.J., Sayavedra-Soto L.A.,
 RA "Characterization of the expression of two distinct alcohol
 RT dehydrogenases involved in butane metabolism in Pseudomonas
 RT butanovora";
 RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF355798; AAK27220.2; -
 DR SEQUENCE 691 AA; 75070 MW; 4FC7FD20CDA14E64 CRC64;

Query Match 14.7%; Score 299.5; DB 2; Length 691;
 Best Local Similarity 24.2%; Pred. No. 2.3e-14;
 Matches 100; Conservative 62; Mismatches 173; Indels 79; Gaps 14;

QY 1 IVANGVIAGSTQYSPGCFVSGHDSATGELMNYFIPRAGEGDDTGNDEYEAAMT 60
 DB 186 VIKGKVLINGGAFFGVSG-YITAYDPTAAS--PRGVVPGGDP-----SLPFDASME 236
 QY 61 GAMGQITDPTNLVHYGSTAVG-----PASGTORGTCGTL 98
 DB 237 AA-AKTWDPAGOVLAGSGRRIGVELDLGKAGCCGTSAPATPSPWHRKRSPPAGDNL 294
 QY 99 GNTNTRAVRPDGTETVWRHOTLPRDNMDQECTFEFMAVTVNDY--QPTSEMGLOSIINPA 156
 DB 295 YTAIVALLRPDGEYVWYHQTPADNMDYTSTODLILADIEIGKRP----- 340
 QY 157 ATGERAVLTGVCCKTGTMMQFPAETGEFLMARDTVYONMTESIDENGLIVTVN-EDAILKE 215

Db 341 ----RKVILHAP-KNGFEVIDRTDGKFIQAQNFVPMNATGDENGRPIENEGAMPGH 395
 QY 216 LDVEYDVCPFFLGGDMPSALNPDSCITFIPLANVCYMAVDOFTSMYNTSN--- 272
 Db 396 LSRF---PAPSKRTMHSISYSPQGLAFPAQNIPL-VLEDKRWMS---YNOAPOQ 447
 QY 273 -----VTKLPKCKDMIGRIDAIDISTGFTLMSVERAANAAYSPVLSTGGVLFN 320
 Db 448 AMAGICMNLGMLVNPBPASQPCGRILAMPVOQKEVWKKEHVSFNGCTLVTAGNVYHQ 507
 QY 321 GGDTRFRFALSOETGTLMOQRLATYASGOAISYEVDGMOYVAILA-GGGSVYS 373
 Db 508 GTADARLLAFDARDKELMSAPMGTCVIAPTYEVDGKQYSAIVAGMGVYGN 561

RESULT 15

09AGW3
 ID 09AGW3 PRELIMINARY: PRT: 623 AA.
 AC 09AGW3:
 DT 01-JUN-2001 (Tremblrel. 17, Created)
 DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE 1-BUTANOL DEHYDROGENASE BOH PRECURSOR.
 OS Pseudomonas butanovora.
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 OC Pseudomonas.
 OX NCBI_TaxID=86174;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Sayavedra-Soto L.A., Vangnai A.S., Arp D.J.;
 RT "Characterization of the expression of two distinct alcohol
 RT dehydrogenases involved in butane metabolism in Pseudomonas
 RT butanovora."
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF26086; AAK15506.1; .
 DR HSSP; Q924J7; 1FEG.
 DR InterPro: IPR002372; Bac_PQQ_repeat.
 DR Pfam: PF01011; Bacterial_PQQ; 4.
 KW Signal.
 FT SIGNAL. 1 29 POTENTIAL.
 FT SEQUENCE 623 AA; 67553 MW; A573A1C64AC12D55 CRC64;

Query Match 13.9%; Score 281.5; DB 2; Length 623;
 Best Local Similarity 24.2%; Pred. No. 4.6e-13;

Matches 103; Conservative 65; Mismatches 161; Indels 103; Gaps 18;

QY 6 VIVAGSTCOYSPFC--FVSGHDSATGSELARNYFI-----PAGEEGDEITNG 51
 Db 199 LTVHGG--GDEFGVGMIFARDPDTGEVYARPMVEGHGRLNGKDSIPTGDKAPSWP 256
 QY 52 NDYEA-----RWMTGA---WGQITDPTNLVHYGSTAVGPASETQRTGP-----GT 96
 Db 257 DDPNSPTGKVEAMSGGAPWOTASFVENNMVYIGANPAPWNTKRTAPGDDPRMWS 316
 QY 97 LYGNTRFAVAPDTEGELVWHRHOTLPRDNMDCECFEMAVTN-----VDYQSTEME 147
 Db 317 LF-TSGQAYVDASTGELKGYOHTPNDAMPSCGNNSVLEFYKDPKTKMNASAHADR 375
 QY 148 GLOSIN-----PNAATGERVLTGVPCKTGTMMQ--FDAETGEFLMARDNTYON 194
 Db 376 GFEFVTRDRMLAKGAGYPNPTS--LIGAMPVVDGITMAGSFDLTKGKI-EKDNRPQ 431
 QY 195 MIESIDENGIVTNEDEILKELDEYDVCTFLGGDMPSALNPDSCITFIPLNNVCYD 254
 Db 432 PKGADGESIFVS-----PPFLGNTMHPMSYSPDTGLFYIPANHAMD 476
 QY 255 MAVDOFTSMYNTSNVT-----KLPG-----KDMIGRIDAIDISTGRTLSVY 301
 Db 477 -----YMTENTYTKAGSAYLGGCFRIKULFDHVGILRAIDPSPARSIGAOG 523
 QY 302 RAAANTSPVLSTGGVLFNGTDYRFRALSOETGTLMOQRLATYASGOAISYEVDGMOY 361

Db 524 RVPA-VAGTLTTAGGWFTGTSDSGLAKAFDAKNGKELMKFOTGSGVVSVPVTWEMDEQY 582
 QY 362 VAIAGGVSYSGSL 375
 Db 583 VAIOSG---YGCAY 593

Search completed: May 24, 2002, 10:17:33
 Job time: 746 sec

Fri May 24 11:27:12 2002

us-08-934-506a-5_copy_180_556.rsp

